Analysis of the Report to the U.S. CPSC by the CHAP on Phthalates and Phthalate Alternatives

(released July 2014)

Prepared for: ExxonMobil Chemical Company
by: ExxonMobil Biomedical Sciences, Inc.

September 2014

EXECUTIVE SUMMARY

ExxonMobil Biomedical Sciences, Inc. (EMBSI) scientists have carefully reviewed the CPSC CHAP report (2014)[1] and its relevance to the anticipated CPSC rulemaking in response to CPSIA section 108[2]. EMBSI expertise includes: human health toxicology, human exposure sciences, epidemiology, biostatistics, and regulatory risk assessment.

We agree that the interim restriction on DIDP should be lifted — even with the excessively conservative methodology employed by the CHAP, this was its recommendation, and the science clearly supports this determination. EMBSI scientists disagree with the CHAP's recommendation to maintain the DINP restriction and the method used to support this recommendation. The CHAP report presents two different risk assessment approaches: individual chemical risk assessment and cumulative risk assessment. This review considers both approaches and offers insight into their quality vs. the state-of-the-science.

Viewed correctly, the information in the CHAP report actually demonstrates that DINP can be used in toys and children's products with a reasonable certainly of no harm to children, pregnant women, or other susceptible individuals with an adequate margin of safety. That is, even without correction of errors in the CHAP report, the CHAP analysis shows that exposures to DINP itself – even in the aggregate of all sources of exposure – are below levels of concern. Furthermore, as demonstrated in the analysis provided here, upon correction of the errors in the CHAP report, it is observed that the margin of safety is yet much greater.

Our detailed analysis reveals that the CHAP report does not reflect the state-of-the-science in chemicals risk assessment and is not of high standard. This is most apparent in their cumulative risk assessment (CRA) conduct – the basis for the CHAPs recommendation to restrict DINP. Analysis of the CRA shows two types of limitations:

- Use of outdated and / or erroneous data, making the assessment outcome irrelevant
- Use of a screening-level assessment, not appropriate for regulatory decision making

However, when these errors are corrected, the updated CRA outcome shows that the cumulative risk from all five phthalates is below a level of concern. Therefore, there is no scientific basis for the CPSC to recommend any restriction on DINP.

In summary, we have identified 7 major concerns with the CHAP report:

1. **Lack of adherence to systematic scientific review process:** The CHAP determined it was not amenable to use the state-of-science methodology for risk assessments as a basis for its

1 Report to the U.S. CPSC by the Chronic Hazard Advisory Panel on Phthalates and Phthalate Alternatives (July 2014) (http://www.cpsc.gov//PageFiles/169876/CHAP-REPORT-FINAL.pdf)

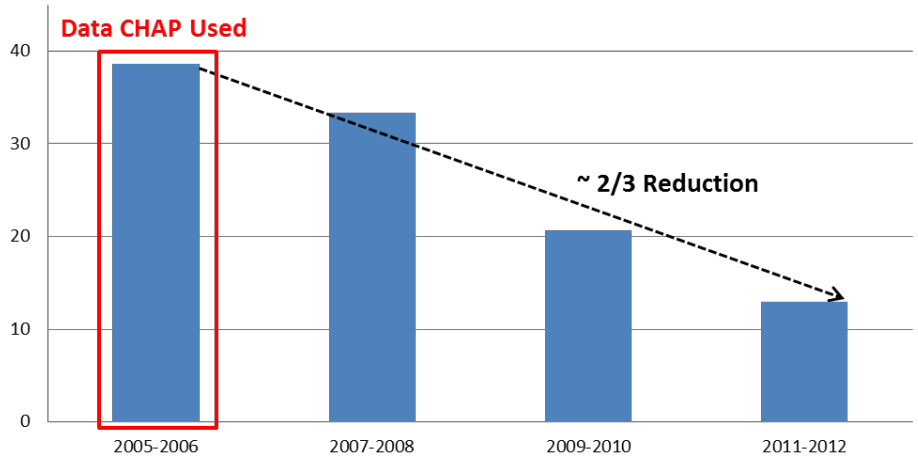
2 Consumer Product Safety Improvement Act of 2008 (http://www.cpsc.gov/PageFiles/129663/cpsia.pdf)

approach[3]. The impact of this determination is that the report is not a meaningful tool that best supports the CPSC with their rulemaking. For example, the CHAP inexplicably ignored applying the established practice of a weight of evidence for data evaluation and instead chose to emphasize only those studies that support its premise that DINP is of public health concern, while not including published studies that did not support the premise[4]. Further, the stated preference for studies that reported effects in preference to those reporting no effects introduces a systematic bias.

2. **Unreliable assessment of toxicity:** All three hazard values (point of departure) used by the CHAP for DINP were based on inappropriate values (*i.e.* outdated, inappropriately modeled data, and /or incorrect) which, in every case, overstated the hazard of DINP. Additionally, though the hazard endpoint chosen for the cumulative risk assessment (endocrine disruption) was assumed relevant by the CHAP, the assumption is highly conservative given the latest experimental data indicating that humans are either less responsive or non-responsive to the endocrine-related effects of phthalates comparison to rats.

3. **Inappropriate human exposure data selected:** The CHAP used outdated NHANES data for use in the risk assessments, despite new datasets being published and available to the CHAP. The most recent NHANES datasets confirms a clear declining exposure trend line, resulting in a >60% reduction in 2011-2012 DEHP urinary metabolite levels over 2005-2006. This means the CHAP has ignored that the potential for cumulative risk from phthalates has dropped by ~2/3. Had the CHAP applied the most recent NHANES data, the cumulative risk outcome would show no risk (i.e., a Hazard Index below 1), confirming that DINP is safe to use in all applications (see Figure 1).

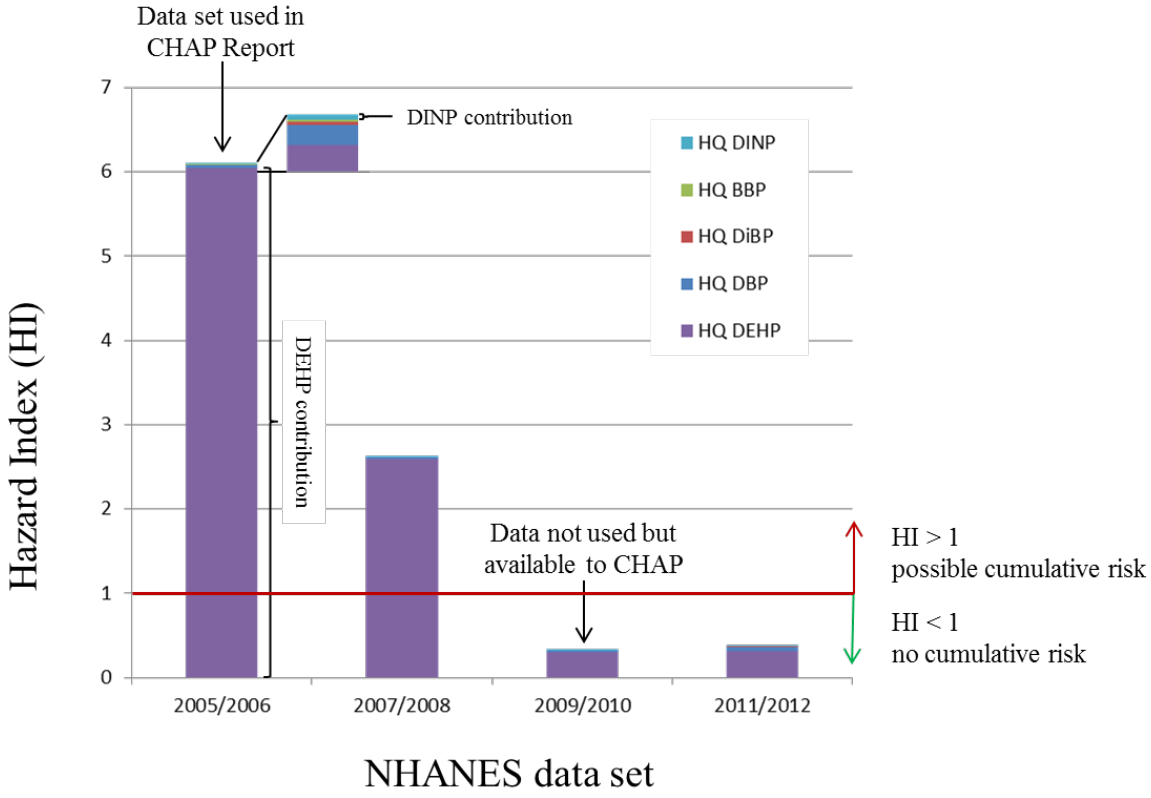
Figure 1 Significant Downward Exposure Trend in DEHP Metabolite Levels Using Mono(2-ethyl-5-carboxypentyl Phthalate (2E/5C) As An Example (95th percentile)[5]



3 Report page 12: '…its review was not amenable to the systematic review methodology.'
4 Report page 21: '…rely on the study reporting adverse effects …'
5 2009-2010 CDC NHANES data published September 2012. The stopping point for CHAP analysis and interpretation was information available by the end of 2012.

We present a re-analysis of the CHAP's approach using the most relevant human exposure data, *i.e.*, the most recent NHANES datasets. This analysis shows that when the CHAP's decision logic is applied to these more relevant datasets, the cumulative assessment would not reach levels of concern for the target population and, as such, there would be no scientific basis to restrict DINP in toys and child care articles. An example for pregnant women is provided below (Figure 2).

Figure 2 A simplified version of Figure 5 pg. 9 of this analysis[6]



4. **Derived inaccurate Margins of Exposure:** Based on the limitations in both the hazard and exposure steps of the risk assessment, scientifically inappropriate Margins of Exposure were identified. Corrected, the Margins of Exposure would be 4-12 times greater, demonstrating far less assumed possible risk to human health.

5. **Unreasonable use of Screening Level Cumulative Risk Assessment as regulatory basis:** It is understood that there is no regulatory agency in the world that regulates industrial chemicals on the outcome of a cumulative risk assessment (CRA). The science simply has not reached a point where such a CRA-based decision could be justified. Further, the CHAP report does not use the state-of-the-science CRA tools (e.g., WHO Framework, 2009), potentially

6 Report page 12: '[t]he stopping point for CHAP analysis and interpretation was information available by the end of 2012.'

reducing the reliability of it CRA. At best, the evaluation was a Screening Level CRA, suited to identify potential risks under "worst case" conditions, rather than a robust cumulative risk evaluation that would form a basis for rational rulemaking.

6. **DEHP and DBP are the cumulative risk drivers, not DINP:** The CHAP report Cumulative Risk Assessment reveals that DEHP and DBP account for ~ 99% of the cumulative risk with DINP contributing <1% to the overall risk estimate. The ~1% worst-case contribution from DINP is less than the range of variation in DEHP's cumulative risk contribution. Nevertheless, this the contribution from DINP is the basis for the CHAP's recommendation to make permanent the interim restriction on the use of DINP.

7. **Failure to consider significance of transition to alternative plasticizers:** The CHAP report notes that the non-phthalate alternative plasticizers included in their assessment are less well researched than DINP.[7] However, they recommend no action by CPSC other than encouraging toxicological testing. The CHAP provided no discussion about the relative risks of substituting a chemical which is well-studied and for which risks are negligible with alternatives that are much less studied.

Conclusion

The CHAP report does not provide the CPSC scientific staff with an accurate assessment of whether an unacceptable risk to toys and childcare articles exists from DINP. This is the result of the CHAP not conducting a high quality risk assessment, not conducting a systematic review of available data or a weight of evidence evaluation, using erroneous data and invalid assumptions, and utilization of outdated exposure data. It is noted that the peer reviewers who provided input to the CHAP in a report finalized August 2013, a year before the release of the CHAP report, highlighted some of these concerns, yet they remain unaddressed in the published CHAP report.

We encourage the CPSC scientists to apply a rigorous assessment of the CHAP report and conduct its own thorough scientific assessment using the most current data available. Fortunately, there is a wealth of data available through the CHAP process and through separately published and more current information published since the last public CHAP meeting in 2012.

7 Report page 14: '*Phthalate Alternatives.* Although data on most phthalate alternatives are limited, …, the CHAP recommends no action at this time.'

Table of Contents

Comparison of the Conclusions of the Current Evaluation of the Safety of DINP in Children's Toys and those of Previous CHAP Conclusions

The CHAP's 2014 report actually provides additional evidence that the prior CPSC conclusion that DINP could be safely used in children's toys (CPSC, 2001; CPSC, 2002; Babich *et al*., 2004) was appropriate and scientifically supported. The previous CHAP identified a liver effect (spongiosis hepatis) as the most sensitive effect of DINP. Using those data, the CHAP selected 12 mg/kg/day as the point of departure and calculated an allowable daily limit (ADI) of 120 µg/kg/day. Estimated background exposures to DINP were well below this limit (as more recently confirmed through urinary metabolite measurements) and estimated exposures of children from toys represented a small incremental amount of the total exposure. The analysis by the CHAP (2001; 2002) coupled with studies of exposure to DINP from toys led CPSC staff to conclude that "oral exposure to DINP from mouthing of soft plastic toys is not likely to present a health hazard to children" (Babich *et al*., 2004).

The more recent CHAP did not identify a more sensitive endpoint to use as a point of departure, and it did not conduct any further investigations of exposure of children to DINP from toys. In short, what the current CHAP process demonstrated was that in the more than 10 years since the initial assessment of safety of DINP in children's toys, there have been no new findings that would change that conclusion. It should be noted that the more recent CHAP did change the focus on the assessment from liver changes to endocrine-mediated effects, but this was not new information. The basic data, including the results of the two-generation reproductive toxicity test (Waterman *et al*., 2000) and the peri-natal study of Gray *et al*. (2000) had been published and were reviewed by the CHAP (2001; 2002) and by the National Toxicology Program Center for the Evaluation of Risks to Human Reproduction (NTP CERHR, Kavlock *et al*., 2002). The more recent data provided quantitative refinement, but there is nothing that is fundamentally at odds with the conclusions of the earlier CHAP (2001; 2002). In this regard it is surprising that the current CHAP (2014) did not discuss why their conclusions differed from those of the previous CHAP.

The CHAP's basis for the recommendation, as detailed below, is derived from faulty data and analysis.

1. The data provided in the CHAP report indicates margins of exposure for DINP that well exceed what is "considered adequate for public health."

2. The CHAP's Cumulative Risk Assessment (CRA) overstates the risk of cumulative exposure by using a sample size that was not sufficient for the analyses conducted, inappropriate points of departure, and outdated exposure data.

3. The use of current exposure data indicates no cumulative risk, the Hazard Index (HI) drops below 1 for all populations. The data was available to the CHAP at the time of the assessment, but not utilized by the CHAP.

In short, the careful review and updated analysis using the approach described in the CHAP report supports that the CPSC can remove the interim ban on DINP with confidence that there will be no unreasonable risk of harm to pregnant women, infants, children or other populations.

Interim Ban on DINP can be Lifted

Adequate Margins of Exposure

The lower bound estimates for the margins of exposure (MOE) estimates reported in the CHAP are not scientifically based. The CHAP calculated a hypothetical no effect level for DINP of 11.5 mg/kg/day (2.3 x 5 mg/kg/day). Use of a hypothetical value as part of the MOE range is not appropriate given well established experimental values already identified in the report. Using 50 mg/kg/day as a conservative estimate, as acknowledged in the CHAP report[8], the lower boundaries on the MOE are approximately 52,000 for pregnant women and > 12,000 for infants.

Table 1 Revision of CHAP Report Table 5.1[9]

Chemical	Range of PoD's (mg/kg-d)	Pregnant Women (NHANES)[10]		Infants (SFF)	
		Daily Intake (μg/kg-d) Median (0.95)	Margin of Exposure PoD/Daily Intake (in same units) Range (0.95)	Daily Intake (μg/kg-d) Median (0.95)	Margin of Exposure PoD/Daily Intake (in same units) Range (0.95)
DINP	**50**-750	1 (11)	**50,000** – 750,000 (**4,500** – 68,000)	4 (18)	**12,500** – 190,000 (**2,800** – 42,000)

Overstated Risk by the CHAP's CRA.

Of note, the cumulative risk assessment employed by the CHAP can be best thought of as a screening process to identify areas of concern (Meek *et al*. 2011, Price *et al*. 2012a). When considered on that basis, the conclusions are: (i) that the group of investigated phthalates does not pose risks to the population at large when the assessments are appropriately, as discussed below, based on median exposure values, and (ii) that any potential concerns relate to situations involving high exposures to DEHP. Three cumulative risk assessments on these phthalates have

[8] "Taking a conservative approach, the CHAP assigns the NOAEL for DINP at 50 mg/kg-day." (pg 98)
[9] Bolded values differ from those in the CHAP report by replacing the inappropriately modeled value for DINP, see discussion on Case 2, with the conservative NOAEL assigned by the CHAP for DINP
[10] Daily intake values for DINP for the more relevant 2009-2010 data set are 2 μg/kg-d (median) and 13 μg/kg-d (0.95) for pregnant women. MOEs using these data are large and protective of public health (3,800-58,000 at the 95th percentile)

been published and all came to similar conclusions (Benson, 2009; Kortenkamp and Faust, 2010; Christensen *et al*., 2014). Because this was a screening assessment, the next steps should have been to examine more critically those situations involving potential risk, *i.e*., situations involving exposures to DEHP at high levels. These next steps were not conducted by the CHAP.

Inadequate Sample Size

An additional issue is the small sample size of pregnant women which constitutes a less robust dataset than "women of child-bearing age". The limited data for pregnant women in the NHANES dataset is highlighted in a communication by CPSC staff with investigator Shanna Swan in which additional data were requested[11]. Given the limited amount of data for this group of individuals, it would seem more scientifically appropriate to use data for women of reproductive age as a surrogate. Another issue relates to the regression models used by the CHAP which account for age, race, gender, and body size in the estimates of daily creatinine excretion (Mage *et al*., 2008). These equations do not fully address fluctuations in creatinine excretion observed during pregnancy (Boeniger *et al*. 1993; Davison and Noble, 1981; Lohsiriwat and Imrittha, 2008). Given these concerns, and the larger datasets for women of reproductive age, this is the population that would have been more appropriate for the HI calculations.

Use of individual data and extreme values as the basis for recommendations is not appropriate. The reasons for this were summarized in a recent report by EPA scientists on the use of urinary metabolites of phthalates, particularly on the use of 95th percentile estimates of exposure in preference to maximum values in cumulative risk assessments:

> "Phthalate metabolites have very short half-lives, on the order of ~5 to 12h (Koch *et al*., 2005; Volkel *et al*., 2002, 2005). Thus urinary concentrations peak shortly after exposure (Kluwe, 1982; Koch *et al*., 2005) and urine sampled during this time of peak concentration could lead to artificially high estimates of daily intake. Conversely, measurements made after concentrations have peaked and declined could lead to artificially low intake estimates…. Although this variability may affect the accuracy of an estimated intake for a single individual, recent work has demonstrated that on the population level, a group of spot urine samples provides a reasonable approximation of concentrations that would have been observed in a population of full-day urine samples collected from the same population for phthalates … Thus, while there may be variability in the tails of the distribution

[11] "They have already applied their method to NHANES data, including women of child-bearing age, but data on pregnant women and infants are lacking." (Chronic Hazard Advisory Panel on Phthalates and Phthalate Alternatives Final Report (2014). Supplemental Materials "SFF Biomonitoring Data Used in Estimating Exposure in the CHAP on Phthalates and Phthalate Alternatives Report (S. Swan)")

(*i.e.*, the extreme highs and lows), the estimated central tendency for the population is likely to be rather stable. Similarly to previous studies of phthalate exposure (Koch *et al*., 2011; LaKind and Naiman, 2011; Marsee *et al*., 2006), we present findings for the 95th percentile of estimated phthalate intake, recognizing that there may be more variability in these values, because this information provides insight into the potential risk at the highest levels of exposure in a general population setting." (Christensen *et al*., 2014)

The variability at the tails of the distribution, in comparison to the central tendency (50th percentile), can be seen in the NHANES data for pregnant women used in the CHAP report (their Cases 1 and 3) by removing the individual with the highest HI value (Fig 3).

Figure 3 Hazard Indices calculated both with, and without the individual with the highest HI value (sample size 130 vs 129) for the 99th, 95th, and 50th percentile.

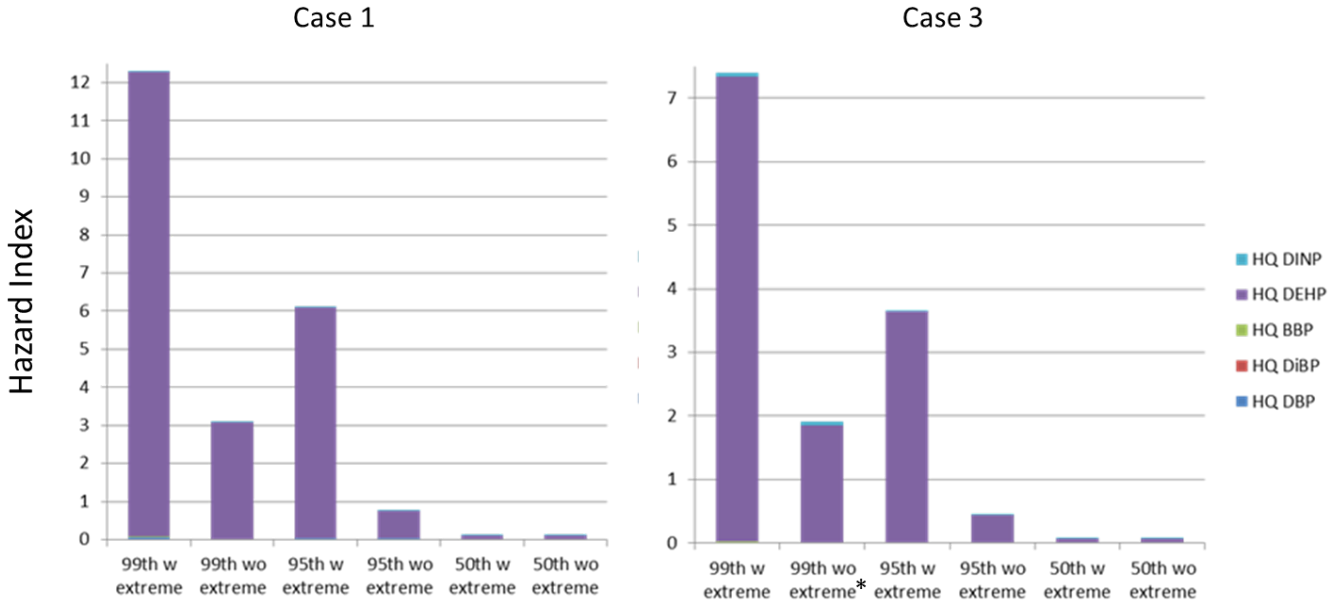


Figure 3 clearly demonstrates the impact of a small sample size and why it is inappropriate to use the tails of the distribution. Removal of one person, the most highly exposed individual, from the analysis has a very large impact on the HI calculation. In both cases the HI calculation at the 95th percentile drops below one. The 95th percentile, in itself, is considered a conservative estimate (as stated in the EPA cumulative risk assessment) and the graphic clearly demonstrates further the degree of instability at the tails for such small sample sizes.

Additionally, in their publication, Kortenkamp and Faust (2010) suggested that the most likely explanation for DEHP at high levels was that these most highly exposed individuals were patients undergoing medical procedures involving phthalate-containing tubing and disposables[12]. Given this is a likely reason why certain individuals are experiencing high exposures, it is

[12] DEHP is the primary phthalate used in medical devices, particularly devices with the most direct human exposure, such as IV tubing.

surprising that the data was used to calculate "worst case" assessments or that at least discussion be provided by the CHAP on confounding factors. It seems paradoxical that exposures to phthalates associated with necessary medical procedures are, in large part, used to justify recommendations to reduce phthalate exposures generally.

The most recent urinary metabolite data provide evidence that exposures to DEHP are declining. As DEHP dominated the cumulative risk assessment, the overall risks of the sum of the phthalates are lower than those calculated by the CHAP, at least in part because of risk management measures that have already been implemented to reduce exposure. The use of current data in the CHAP methodology indicates no basis to extend the interim ban on DINP to provide reasonable certainty of no harm.

Inappropriate Points of Departure

The CHAP defined 3 cases, each of which was based on a different point of departure for DINP. In each case the point of departure was incorrectly determined, resulting in exaggerated estimates of the potential risks associated with exposure to DINP.

Case 1 (CHAP, 2014, p. 64) was based on a cumulative risk assessment published by Kortenkamp and Faust (2010). The point of departure for DINP was based on a study (Gray *et al*., 2000) in which it had been reported that daily exposure to 750 mg/kg DINP resulted in a statistically significant increase in retained nipples in rats. Kortenkamp and Faust then used an assessment factor of 500 (presumably factors of 10 each for interspecies and intraspecies sensitivity and a factor of 5 because the analysis was based on a lowest observed effect level). However, during the period that the CHAP was conducting its assessment, Boberg *et al*. (2011) published a dose response study of DINP in which it was shown that 750 mg/kg was an effect level for nipple retention, and that the no observed effect level was 600 mg/kg/day. It is always more scientifically correct to use a no effect level when available than a low effect level for assessment purposes. Had the correct value been used, the point of departure for DINP in case 1 would have been 6000 μg/kg/day (600000/100) rather than 1500 μg/kg/day (750000/500). In this case the risk attributed to DINP was overestimated by four-fold or greater. We understand that in case 1 the CHAP followed Kortenkamp and Faust (2010), but as the assumptions were shown to be incorrect in subsequent publications that the CHAP reviewed, it is scientifically inappropriate for them to have continued to use this value in an assessment charged to reflect the most relevant information.

Case 2 was based on a study by Hannas *et al*. (2011) which compared the effects of phthalates including DEHP and DINP on testosterone production from male rat fetuses under *in vitro* conditions. The CHAP noted that, in this assay, DEHP was 2.3-fold more active than DINP. The CHAP then hypothesized that the same relative relationship would hold for other endocrine-related effects. Using the no effect level for DEHP (5 mg/kg/day based on reproductive tract malformations, delayed vaginal opening and decreased sperm production), the CHAP calculated a hypothetical no effect level for DINP of 11.5 mg/kg/day (2.3 x 5 mg/kg/day) and used that value as the point of departure for case 2. However, had the CHAP compared their theoretical values with the empirical evidence, it would have been determined that their model was

unsupported. In fact, this specific issue was discussed in a publication by Gray *et al*. (2000) in which it was estimated that DINP was 10-20 fold less potent than DEHP. In short, the point of departure for DINP used in case 2 was based on a hypothesis that could have been easily tested and shown to be unsupported. Based on the estimates of a 10-20 fold difference by Gray *et al* (2000), the risk attributable to DINP were overestimated by factors ranging from 4.3 to 8.6.

It should be noted that the NOEL for reduced testicular testosterone production, supported by Hanna *et al*. (2011) and Clewell *et al*. (2012a), is at least 100 mg/kg-day. Reduced testosterone is considered an early biomarker of the potential for testicular tract malformations at higher doses. It is not scientifically supportable that a hypothetical NOEL (11.5 mg/kg/day) for these effects could be lower than that observed for the biomarker (100 mg/kg/day) in an absence of the effects. This information clearly demonstrates the underlying assumptions and hypothesis for the model are incorrect.

Case 3 uses a point of departure for DINP which, as cited, is an error. As listed in Table 2.1 (CHAP, 2014, p. 24), the NOAEL that the CHAP used for DINP was 50 mg/kg/day based on increased nipple retention, referencing Boberg *et al*. (2011). However, Boberg *et al*. actually reported that the incidence of retained nipples was significantly elevated at 750 mg/kg/day but that the difference at 600 mg/kg/day was not significantly different from control values. In short, the correct value is not 50 mg/kg/day but 600 mg/kg/day, and risk for DINP was overestimated by a factor of 12. A value of 50 mg/kg/d is the NOEL for testosterone reduction in Clewell *et al* 2012a, which in context with Hannas *et al*. 2011 puts the true NOEL for testosterone reduction at least 100 mg/kg/d. Therefore, correct usage of data in this case indicates the risk attributable to DINP was overestimated by at least a factor of 2.

The comments above were specific to DINP. However, it should also be noted that different points of departure were used for other phthalates. On page 62 it is stated that "case 3 includes values from the CHAP's *de novo* literature review of reproductive and developmental endpoints focused on reliable NOAELS and PDOs (table 2.1, CHAP, 2014, p. 24)" (emphasis added). In Table 2.1, the NOAELs for DBP and BBP are given as 50 mg/kg/day and that for DIBP is 125 mg/kg/day, all values based on references given in the table. These were converted to potency estimates for antiandrogenicity (PEAA) values through the use of a 100 fold safety factor. The values in case 1 were taken from Kortenkamp and Faust (2010) in which the points of departure for these phthalates are given as 20 mg/kg/day, 40 mg/kg/day, and 66 mg/kg/day respectively, and each of these values was divided by an uncertainty factor of 200 due to the limited size of the study. In round numbers, what this means is that the points of departure for DBP, BBP and DIBP that were used for case 1 were approximately four-fold lower than the values used in case 3. The value used in case 1 for DEHP was 3 mg/kg, whereas the value in Table 2.1 is 5 mg/kg/day. As all the PEAA values for all of the phthalates evaluated differed from those that the CHAP considered reliable, the case 1 analysis is misleading.

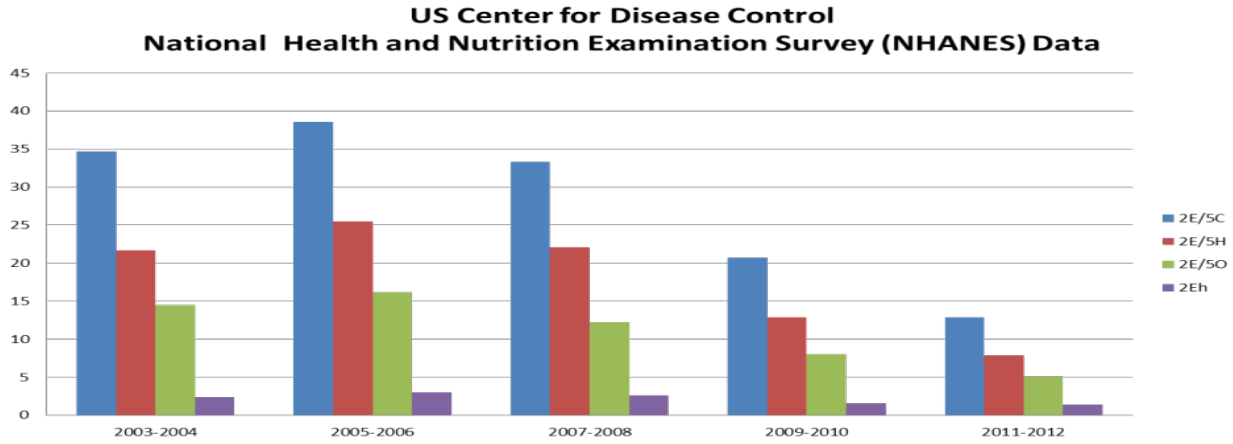
Outdate Exposure Data was Used

As indicated above the CHAP did not use the most recent exposure data for phthalates. The use of outdated exposure data is noted as a limitation on pg 41 of Appendix D (of the CHAP report):

> "A limitation of the analyses presented here is the use of exposure data from 2005–06 for NHANES and 1999–2005 for the SFF. Since these data were collected, the Consumer Product Safety Improvement Act restricted some of the uses of the five phthalates evaluated. The impact on exposure is unknown and not accounted for in the calculation of the HI."

Given that DEHP "dominated"[13] the HI value and was one of the phthalates that is restricted, an evaluation of DEHP metabolite levels in more recent NHANES data sets would give an idea of the potential impact. In addition to the noted limitation, reduction in use of DEHP should be expected not only related to the CPSIA ban but also to the phase out under REACH and other factors that have moved the market away from DEHP. A clearly decreasing trend can be seen in exposure to DEHP metabolites in the later sets of NHANES data (Fig 4). The 2009/2010 data was available to the CHAP but not utilized in their assessment.

Figure 4 DEHP Urinary Metabolite data from NHANES



Geometric means of urine concentrations (μg/L) for the U.S. population from the National Health and Nutrition Examination Survey for four DEHP metabolites (Mono-(2-ethyl-5-carboxypentyl) phthalate (2E/5C), Mono-(2-ethyl-5-hydroxyhexyl) phthalate (2E/5H), Mono-(2-ethyl-5-oxohexyl) phthalate (2E/5O), and Mono-2-ethylhexyl phthalate (2Eh)).

As is shown in Figure 4, the urinary concentrations of DEHP have declined since the 2005/2006 survey, and, by the 2011/2012 survey were approximately 1/3 the levels measured in the 2005/2006 survey. Of note, since DEHP accounted for approximately 90% of the risk estimated with the cumulative risk assessment, the potential for risk calculated from the urinary metabolite information given in the 2005/2006 NHANES report is substantially over-estimated with respect to the current situation. Substantial reduction in urinary metabolite levels were evident in the

[13] "In all three cases studied, the HI value was dominated by DEHP because it had both high exposure and a low PEAA." Appendix D – 40 CHAP report

2009/2010 report, posted to the NHANES website in September of 2012. These data as well as those in the 2007/2008 report were publicly available within the time frame set by the CHAP for data inclusion and were approximately half those reported in 2005/2006. Given the availability of the 2007/2008 and 2009/2010 data sets, it is hard to see the CHAP position quoted above as anything but evidence of the superficial nature in their investigation. In fact, it appears that regulatory actions resulting in restrictions of some uses of phthalates have had an impact on exposure, and this is clear on the basis of urinary metabolite studies that were publically available during the time the CHAP was preparing its report. Had these data been used, the calculated hazard indices would have been substantially lower.

No Risk Found Using Relevant Exposure Data from Cumulative Risk Assessment

To assess the impact of the changes in urinary metabolites on the risk assessment process, we re-calculated the hazard indices (HI) following the CHAP methodology, including the selected points of departure (PEAAs in CHAP report) but using urinary metabolite data from the NHANES 2009/2010 report. The CHAP methodology was followed exactly to clearly depict the effect of using the recent exposure data. In addition to calculating HIs for pregnant women as was done by CHAP, HIs were also calculated for women of child bearing age (15-45)[14] as this latter group of individuals provides a much larger sample size.

Finally, the re-calculations used the 95th percentile values as the upper limits of population exposure for the reason discussed previously. Note, however, that the points of departure used in these re-calculations were the same values as those used by CHAP in its original calculations.

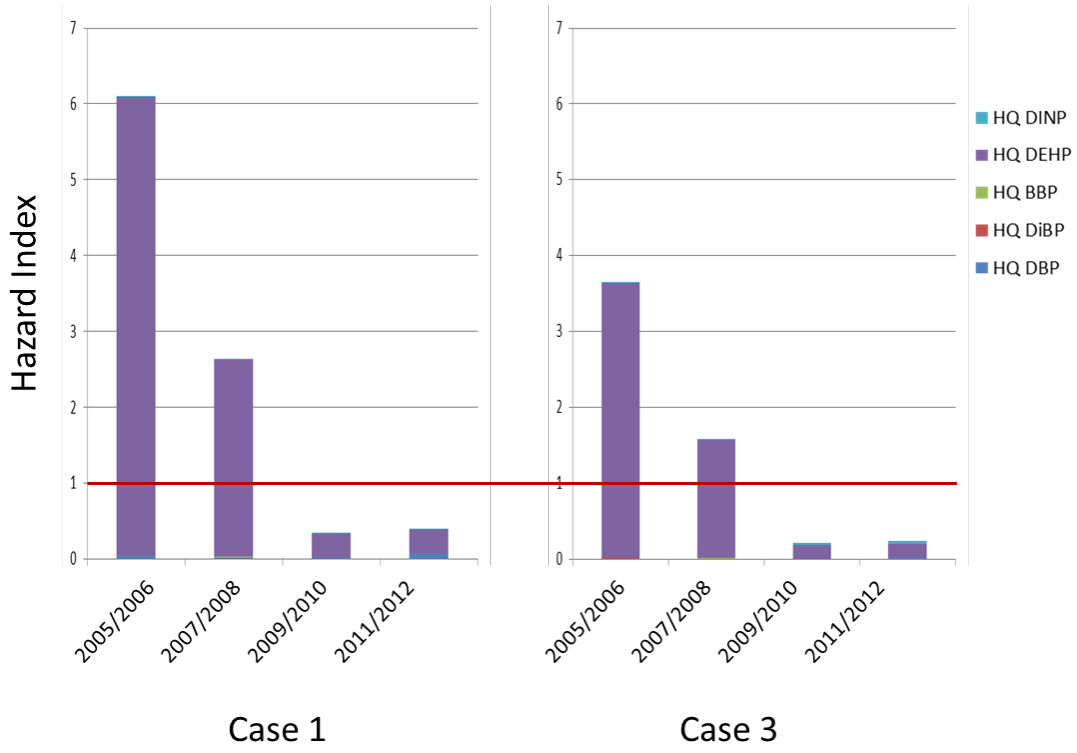
To be clear, the data on pregnant women from each NHANES dataset is given to draw a parallel between the data presented in the CHAP report and how that data would look using the recent exposure data, not as a basis for recommendation.

Figure 5a demonstrates a clear reduction in HI over the NHANES sampling years. By 2009/2010 there is no cumulative risk for the five phthalates evaluated. The same trend can be seen for women of reproductive age (Fig 5b). Note that “Case 1” and “Case 3” used the points of departure for DINP as defined in CHAP Table 2.15 (CHAP, 2014, p. 66). We considered that the point of departure used for “Case 2” was not technically supportable and it is not used in these illustrations.

[14] As defined in the CHAP report.

Figure 5[15] Hazard Index calculated based on the 95th percentile urinary metabolite data taken from the NHANES 2005-2006 through NHANES 2011-2012 reports for (a) pregnant women and (b) women of reproductive age (15-45). NHANES values include sampling weights and methodology for calculating Daily Intake and Hazard Quotients (HQ), as laid out in the CHAP report, were followed.

(a)



(b)

[15] Though aspects of all Cases have flaws, as discussed above, Case 2 was not modeled given the underlying assumptions of Case 2 are inconsistent with the scientific evidence.

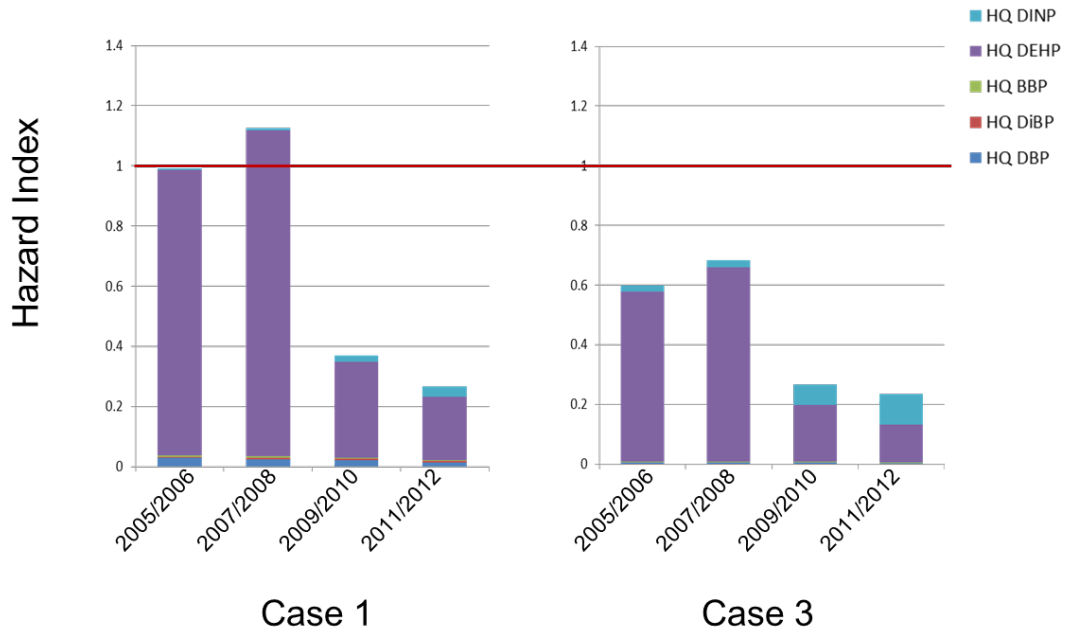


Figure 5 illustrates 2 points:

First, by comparing the upper and lower panels, the hazard indices are substantially higher when based on urinary metabolite data from pregnant women than when based on urinary metabolite data for women of child-bearing age. It is not clear to what extent this is simply a consequence of sampling size and sampling strategy, but the NHANES data represent a statistically-based representative sampling of the women aged 15-45 in the US population, whereas the pregnant women were a much smaller group.

Second, and more importantly, the figures all illustrate the impact of declining levels of urinary metabolites of DEHP. In all cases, DEHP represents by far the largest fraction of the attributed risk, so the use of the more recent NHANES information results in overall hazard index values that are far below 1, the nominal level of concern. Nevertheless, even with the very large declines in urinary metabolites that were documented in the 2011/2012 NHANES report, DEHP still dominates the overall risk estimates. As shown in this illustration, the potential risk identified in the cumulative risk assessment based on urinary metabolite data for pregnant women and the borderline risk identified in women of reproductive age based on the NHANES 2005/2006 report had been substantially reduced by 2009-2010, and were even further reduced when the 2011/2012 urinary metabolite data are used in the CRA, as shown in Figure 5[16]. This re-analysis demonstrates the use of the 2005/2006 data in calculating the HI values results in

[16] The CDC removed the 2011/2012 phthalate metabolite data after it was retrieved for the analysis shown here, due to NHANES errors in sample weights. These errors could affect the depicted 95th percentile, however, the underlying exposure data and HIs calculated should be accurate since they are not calculated with the weighting. The potential difference would come from how individual values extrapolate to the greater population.

substantial over-estimates of risk. Clearly, the recommendation to make permanent the interim ban on DINP, based on the CRA presented in the CHAP report, is not supported by the most recent exposure data. Conclusions that can be drawn from the CRA using the recent data are 1) there are no areas of concern as all hazard index values are < 1, 2) DEHP remains the primary driver of any cumulative assessment, and 3) the removal of the temporary ban on DINP would not alter either 1 or 2.

Detailed Comments on the CHAP Report

Provided below in sections 1-10 are more detailed comments on the entirety of the CHAP report. These comments are based, in part, on the understanding outlined below of the CHAP objectives, requirements of the CPSIA, CHAP recommendations, and the overall process.

Comments Related to the CHAP Objectives

In 2008, the Consumer Product Safety Improvement Act (CPSIA) directed the U.S. Consumer Product Safety Commission (CPSC) to convene a Chronic Hazard Advisory Panel (CHAP) "to study the effects of all phthalates and phthalate alternatives as used in children's toys and child care articles." The CPSIA listed a number of specific objectives for the CHAP, one of which was to provide recommendations to the Commission relating to the use of certain phthalates and phthalate alternatives. The CPSIA also directed the Commission to determine whether or not to continue the interim prohibition of diisononyl phthalate (DINP), diisodecylphthalate (DIDP), and diisooctyl phthalate (DNOP) in children's toys. In discharging this requirement the Commission is to determine whether the ban was necessary "*to ensure a reasonable certainty of no harm to children, pregnant women or other susceptible individuals with an adequate margin of safety.*" This statutory requirement echoes the technical directions to the CHAP, specifically to "*consider the level at which there is reasonable certainty of no harm to children, pregnant women or other susceptible individuals and their offspring considering the best available science and using sufficient safety factors to account for uncertainties regarding exposure and susceptibility of children, pregnant women, and other potentially susceptible individuals.*"

In its Report, the CHAP identified the potential for endocrine-mediated processes as the critical effects and women of child-bearing age as the most sensitive population (since male reproductive development occurs primarily *in utero*), and estimated exposure using urinary metabolite information primarily from NHANES datasets. The CHAP developed a metric, the "potency estimate for antiandrogenicity (PEAA)" as a comparative term for use in the calculation of hazard indices. In the publication on which this analysis was based (Kortenkamp and Faust, 2010), the "points of departure" (the lowest effect or highest no effect levels from animal studies) were divided by uncertainty factors to produce reference doses (RfD, oral). Kortenkamp and Faust (2010) considered that hazard index ratios > 1 were a basis for concern. The CHAP (2014) used the same points of departure and uncertainty factors for their "case 1" assessments as those used previously by Kortenkamp and Faust, but did not use the term RfD which they considered to be a regulatory value. However, unlike Kortenkamp and Faust (2010), the CHAP (2014) did not define decision points or margins of exposure that they considered to

be safe. This omission to define a decision process had implications for the CHAP recommendations as discussed in more detail below.

With respect to those phthalates covered by the interim ban, the CHAP found no evidence that either DIDP or DNOP could affect male reproductive development and appropriately recommended that the interim bans for these substances be lifted. For DINP, the CHAP considered that there was the potential for male reproductive effects based on evidence of reduced testosterone synthesis, but, for all of the analyses conducted, the margins of exposure for DINP were greater than 100. This confirmed the conclusions of the previous CHAP (CPSC, 2001; CPSC, 2002; Babich *et al*., 2004) that DINP could be safely used in children's toys. Further, a cumulative assessment showed that DINP made essentially no contribution to the overall potential risk of the group of phthalates that the CHAP assessed.[17] Thus, in fact the CHAP assessment, in fact demonstrated that, to a reasonable degree of certainty, DINP presents no harm to any segment of the population, either via toys or otherwise, which is contrary to their recommendation that the ban on the use of DINP in children's toys be made permanent. This recommendation seems inconsistent with the CPSIA charge elements for the CHAP, including risk assessment considerations, and is not aligned with the scientific assessment. As the CHAP's risk assessment (particularly if recent exposure data were utilized by the CHAP) demonstrates that there is a reasonable certainty of no harm from use of DINP in children's products, and that there will be no unreasonable risk of harm to pregnant women, infants, children or other populations.

Comments Related to the Requirements of the CPSIA

With respect to the specific CPSIA requirements for the CHAP, the following general comments are provided.

As stated in the CHAP report, the CPSIA requires the CHAP to:

"*complete an examination of the full range of phthalates that are used in products for children and shall –*

(1) Examine all of the potential health effects (including endocrine disrupting effects) of the full range of phthalates – At the beginning of its deliberations, the CHAP made a decision to define the critical effects of the phthalates of interest as those associated with reductions in androgen synthesis. Throughout the public meetings, the CHAP focused nearly exclusively on endocrine mediated effects. In fact, the majority of the toxicological information for the DINP assessment is found in the summary produced by the CPSC staff and was evaluated in the previous CHAP process (CPSC, 2001; CPSC, 2002). It is not clear that all of the potential health effects for all phthalates were critically examined by the CHAP.

(2) Consider the potential health effects of each of these phthalates, both in isolation and in combination with other phthalates – The focus on endocrine-related endpoints limited the

[17] Di(2-ethylhexyl) phthalate (DEHP), di-n-butyl phthalate (DBP), butylbenzyl phthalate (BBP), and diisobutyl phthalate (DIBP) along with diisononyl phthalate (DINP).

evaluation to those phthalates for which there was evidence of endocrine-related effects which, according to the CHAP, were DEHP, DBP, BBP, DIBP, and DINP. This decision restricted the extent of the evaluation as some phthalates are not known to cause endocrine-related effects, and there was little relevant data for the phthalate alternatives.

(3) *Examine the <u>likely</u> levels of children's, pregnant women's, and others' exposure to phthalates based on a <u>reasonable</u> estimation of normal and foreseeable use and abuse of such products* (emphasis added) – We agree with the decision to use urinary metabolite data, particularly the use of NHANES data, for exposure estimation. However, "likely" and "reasonable" are likely best represented by the mean and 95th percentile values which, at least with respect to the NHANES data are based on a large and representative population. The use of data from maximally exposed individuals is not consistent with the CPSIA requirements that the exposure estimates be likely or reasonable, nor is this approach consistent with normal scientific practice as demonstrated in the CRA on phthalates conducted by EPA scientists (Christensen *et al.*, 2014). Analyses based on the extreme values are not reliable.

(4) *Consider the cumulative effect of total exposure to phthalates, both from children's products and from other sources such as personal care products* – It is to be noted that this requirement cannot be directly addressed through the use of urinary metabolite information which integrates exposure from all sources, but rather requires an indirect approach. The CHAP's analysis using empirical data shows that children's products are a minor source of exposure and, thus, would contribute very little to the already very low risk posed by DINP aggregate exposures. The conclusions of the previous CHAP, *i.e.*, that DINP is safe for use in children's toys (CPSC, 2001; CPSC, 2002), is further endorsed upon analysis of the 2014 CHAP Report.

(5) *Review all relevant data, including the most recent, best-available, peer-reviewed scientific studies of these phthalates and phthalate alternatives that employ objective data collection practices or employ other objective methods* – Because the CHAP identified endocrine-mediated effects as the most critical, there were only a few publications that appear to have been considered, and, as discussed in more detail below, it is not clear how carefully these were reviewed. As for non-phthalate alternatives, there was little information available and the CHAP was unable to provide any recommendations other than to obtain more toxicological information. Finally, with respect to exposure, the CHAP used the 2005/2006 urinary metabolite data from NHANES, but the more recent NHANES data (2007/2008; 2009/2010; 2011; 2012) provide clear and objective evidence of declining levels of urinary metabolites of DEHP (the major driver of risk in the cumulative assessment). The use of this more recent information would have resulted in substantially lower levels of estimated risk.

(6) *Consider the health effects of phthalates not only from ingestion but also as a result of dermal, hand-to-mouth, or other exposure* – As indicated previously, exposure estimation based on urinary metabolites integrates exposure from all routes. With respect to empirical data, for phthalates such as DINP and DIDP, percutaneous absorption is very

low and the contribution to total exposure from dermal contact can probably be ignored, though they were included in the exposure estimates by the CHAP.

(7) *Consider the level at which there is reasonable certainty of no harm to children, pregnant women or other susceptible individuals and their offspring considering the best available science and using sufficient safety factors to account for uncertainties regarding exposure and susceptibility of children, pregnant women, and other potentially susceptible individuals* (emphasis added) – The CHAP used assessment factors of 100-500 (Table 2.1.5, p. 66) in their calculations of margins of exposure (MOE) (Table 5.1, p 80), but the CHAP did not provide decision criteria, so "sufficient safety factors" and "reasonable certainty" were undefined. Thus the CPSIA charge appears to not have been met. However, more importantly, the CHAP recommendation that the temporary ban on DINP in children's toys be made permanent appears to be inconsistent with their own risk assessment process (again, particularly considering recent exposure data was not utilized). The CHAP's assessment shows no unacceptable risk from DINP itself. The use of the CHAP screening level cumulative risk assessment is unreliable to make decisions for several reasons explained herein, but at any event, the contribution of DINP in children's products to that cumulative risk is negligible.

(8) *Consider possible similar health effects of phthalate alternatives used in children's toys and child care articles* – The CHAP did not make any recommendations relating to phthalate alternatives, other than to obtain additional toxicological information on them. At least in part, this is because the alternatives have not been as extensively studied and critical information is not available. This exposes a weakness of the evaluation process. Unless there is an agreement on data necessary to qualify a substance for use in children's toys and an evaluation process is defined, it is not possible to draw science-based conclusions about alternatives other than perhaps that they are different.

Comments Related to the CHAP Recommendations

The CPSIA directed the CHAP to undertake a comprehensive examination of phthalates and phthalate alternatives and to determine whether any (other than the already permanently banned DBP, BBP and DEHP) should be declared "banned hazardous substances." For its examination, the CHAP identified endocrine-related effects as the critical endpoints, and, based on these data conducted a cumulative risk assessment which assessed the overall risk of DBP, BBP, DEHP, DIBP and DINP. The CHAP found no evidence that DMP, DEP, DNOP and DIDP had endocrine modulating properties, and these phthalates were not included in the cumulative risk assessment. The alternatives were considered separately but as the data were limited, phthalate alternatives were not assessed.

Taking the screening level cumulative risk assessment at face value, the risks presented by this group of phthalates to the population at large are relatively low; when the analysis was based on median exposure estimates, the hazard index values for the three cases (see more detail below) were approximately 0.2, well below the concern level of 1. To the extent that concerns were identified, these were confined to the most highly exposed segments of the population and associated almost exclusively with DEHP. More specifically, of the fractional hazard quotients

comprising the hazard index, approximately 90% was associated with DEHP, 10% of the risk was associated with DBP, and the risks associated with the remaining phthalates (DIBP, BBP and DINP) were negligible. In the more extreme cases, the attributed risk is almost entirely associated with DEHP. These conclusions are in line with, but quantitatively different from, three previously published risk assessments (Benson, 2009; Kortenkamp and Faust, 2010; Christensen *et al*., 2014).

The CHAP recommended that no action be taken on the permanent bans of DBP, BBP and DEHP; that the interim ban on DINP be made permanent; and that the interim bans on DNOP and DIDP be lifted as there was no evidence that these substances affected endocrine-related properties. Some of these recommendations are consistent with the CHAP's scientific evaluation but others are not. As the assessment showed that there might be risks from exposure to DEHP and DBP, particularly to the most highly exposed individuals, the recommendation to take no action on (*i.e.*, to leave in place) the permanent bans on DEHP and DBP might be appropriate based on their scientific evaluation. However, the analysis provided no reason to continue the ban on BBP. With respect to DINP, the analysis showed that there were wide margins between exposures and potential effects. As indicated, the CHAP analysis demonstrates the contribution of DINP to the cumulative risk is negligible. Nevertheless, the CHAP recommended that the interim ban be made permanent because of that negligible potential contribution. Further, as discussed in more detail below, the cumulative risk methodology was screening level only, used outdated exposure information, and relied on an excessively conservative hazard assessment for DINP. When these factors are taken into account, there is no scientific basis for finding DINP in children's products would pose an unreasonable risk.

General Comments Related to the Overall Process

The principal concern with the CHAP report was that, at least with respect to DINP, the recommendation is not scientifically justified. The data and analysis of the CHAP report, when reviewed critically, show that the CPSC's previous conclusion was correct and that there is no basis to continue the interim ban on DINP. However, a second concern was that the CHAP was unable to assess the potential hazards of phthalate alternatives. In some respects this, was a consequence of the process that the CHAP adopted: the decision by the CHAP to focus on endocrine-mediated effects as the endpoints of concern resulted in a process that was both data intensive and data specific. The CHAP did not have enough information on the alternatives to conduct at least the same type of analysis as it had for the phthalates of interest, and, in the end was unable to make any recommendations other than to obtain more data. This raises questions about the internal processes on which the CPSC relies. Clearly the CPSC needs to have a clear and consistent evaluation process that identifies the data that would be required, rather than addressing issues of constituent safety on a retrospective basis. This would likely help the toy manufacturers assess the safety of the substances they use and the specific data that they would need.

Conclusions

The CPSIA required that a CHAP be convened to study the effects on children's health of phthalates and phthalate alternatives used in children's toys and childcare articles and to make

recommendations regarding those phthalates and alternatives that should be declared "banned hazardous substances." The CHAP adopted a screening risk assessment process that validated a determination by a previous CHAP that DINP can be used safely in children's toys. Even with this screening level, conservative cumulative risk assessment process, for DINP, Margins of Exposure were all greater than 100, deemed adequate for public health, and the HQ had a negligible contribution to the cumulative risk. Of interest, the standard error for DEHP's HQ is greater than the point estimate calculated for DINP. Nevertheless, the current CHAP recommended that the temporary ban on DINP be made permanent. This recommendation reduces the confidence in the science-based approach to assessing product safety, as the CHAP report indicates there is in fact no scientific basis for extending the interim ban on DINP.

In addition to the points summarized above, we have provided more detailed comments relevant to specific pages and details within the CHAP report in which we show that there were several errors in the current cumulative risk assessment that exaggerated the potential risk of DINP. In addition, there are questions about the exposure assessment, and underlying science issues related to the reliance by the CHAP on extreme values. There is also more recent information, not considered by the CHAP, that indicates that regulatory actions already taken likely have addressed concerns related to the use of other phthalates. This calls into question whether any further regulatory actions can be scientifically justified.

1. Issues Related to Cumulative Risk Assessment

1.1 Improper Use of a Screening Level Assessment

The methodology used by the CHAP applied worst-case or high-end assumptions in the assessment of risk. This approach is very conservative and consistent with a screening level risk assessment in which health protective assumptions are appropriately used for parameters employed in calculating exposures and hazards to assure that potential risks are not underestimated. However, screening level assessments such as these are not designed to provide precise estimates of risk; rather they are a means of quickly identifying areas of potential concerns. When a screening level assessment indicates an acceptable level of risk, the assessor has a high degree of confidence that the potential risks are much lower than the calculation and, therefore, the true risks are lower and/or perhaps non-existent. However, when a screening level risk assessment indicates a potential concern for a health or environmental effect, this does not mean that the true risks are significant and warrant action. Rather, it means that the risk evaluation should be refined using more realistic and accurate parameters in the methodologies to calculate risks. The outcome is then a refined risk assessment that more accurately quantifies actual risks. Therefore, while results from the screening level approach that the CHAP used is an appropriate method to identify situations of possible concern, it is not an appropriate basis for further regulatory action without further refinement. For example, one possible next step would be to utilize the most relevant exposure information. As described in these comments, utilization of the most relevant exposure information, even in the absence of further refinements to the risk assessment approach, further indicates that potential risk of exposure to DINP is negligible.

1.2 The Range of Results was not Properly Characterized

The CHAP did not provide any estimate of the magnitude of the uncertainty in the identified potential risks that may have resulted from the assumptions applied throughout the assessment. The full quantitative impact of using a different point of departure, different uncertainty factors for quantifying the toxicity value, or a different percentile when quantifying exposure, is not clearly presented in the assessment or appendices. While it is recognized that the CHAP does not want to under-estimate risks, using a conservative individualized approach in the exposure estimate, in conjunction with other conservative decisions in the assessment, undermines the value of the assessment, and, as documented above, resulted in a CRA that is unrealistic. The representation and communication of concerns (in this case Hazard Indices) in an assessment intended to inform regulatory decisions should more accurately reflect scientific uncertainties, include assessment of the sensitivity of derived estimates to model assumptions and end points selected, and employ appropriate tabular and graphic displays to illustrate the range of the estimates and the effect of uncertainty of the estimates.

In a number of places throughout the document the CHAP mentions three different cases to "*determine the sensitivity of the results to the assumptions for PEAAs and the total impact on the HI approach*"(pp 4, 62, D-19, etc). While this approach is helpful for understanding the impact the point of departure selection may have had on the potency estimate of anti-androgenicity, the cases fall short of thoroughly quantifying the impact as they fail to provide a range of results for each case. More specifically, as uncertainty is inherent to each case study and therefore assumptions applied, the potential impacts of these uncertainties should be fully quantified and transparently reported. Only then is the impact of the assumptions in each case study fully apparent. The CHAP seems to be aware of the limited scope of the performed sensitivity analysis as reflected in the following statement in section 4.1. "[a] second case for evaluating the HI was undertaken so that the sensitivity of the results to <u>some</u> of the underlying assumptions could be assessed." However, the CHAP fails to highlight or even qualitatively assess the impact the ignored assumptions could have on the assessment including the selected uncertainty factors, appropriateness of the model, and statistical variation within the datasets. Again, the cumulative risk assessment that was used by the CHAP was only a screening level assessment and was not sufficiently scientifically supportable to serve as the basis for regulatory action.

1.3 Uncertainty was Poorly Characterized

While the CHAP report does acknowledge uncertainty throughout the assessment, they do not indicate how these uncertainties impact the outcome of their assessment, nor do they adequately describe how these uncertainties were considered in their conclusion. The authors used words such as "probably" "more or less", and "not entirely consistent" when describing how the data supported their conclusions throughout the report. Use of these terms indicates uncertainties with the assessments, but it is not transparent nor is it objective with respect to how these uncertainties were integrated, and gives the reader the impression that the assessment is

impartially biased toward high-end conservative estimates. While this is acceptable for a screening level assessment, the CPSC should evaluate and present the full range of impacts of these assumptions on the outcome transparently and objectively. Presenting this information transparently will allow for a complete understanding of the range of risks identified and the impacts major assumptions have on the identified risks. A complete understanding of the impact of all major assumptions is critical to making informed and science-based decisions.

1.4 Underlying Assumptions in the Cumulative Risk Assessment

1.4.1 Assumptions Inherent in the Cumulative Risk Assessment Should be Scrutinized.

Assumption 1: A cumulative risk assessment could be conducted on a group of phthalates as indicated in the CPSIA based on evidence of their ability to similarly disrupt male sexual differentiation in reproductive toxicity models in rats (*i.e.*, exhibited effects characteristic of the androgen insufficiency syndrome).

At approximately the same time as the passage of the CPSIA, the National Academy of Sciences (NAS, 2008) recommended that "Accordingly, the cumulative risk assessment of phthalates should consider any chemical that leads to disturbance of androgen action and is thus capable of inducing any of the effects on the development of the male reproductive system that are characteristic of phthalate exposure." The statement implies that disturbance of androgen action indicates capacity for inducing hypospadias, cryptorchidism, reproductive tract malformations, a decrease in Leydig cell function, a decrease in AGD and or a decrease in fertility.

The weight of the evidence indicates that in rats, exposure to DINP during the later stages of fetal development results in a transient reduction in testosterone production (Clewell *et al.*, 2012b). However, effects on the development of the male reproductive system are not observed at doses below the commonly accepted limit dose of 1000-2000 mg/kg/day (Waterman *et al.*, 1999). Therefore, a number of non-overlapping disrupted pathways may result in the varied and complex responses induced by certain phthalates. This complexity highlights the need to carefully examine the specific toxicity, adverse health effects, and potentially associated events for each individual phthalate. Given the differences in toxicological effects between the different phthalates, it is plausible that multiple modes of action may be at play; observation of a single precursor event (e.g., reduced testosterone) may not be predictive of the capacity for "inducing any of the effects on the development of the male reproductive system that are characteristic of phthalate exposure."

The CRA used anti-androgenicity as the endpoint. It is questionable that the anti-androgenic effects seen in studies of DINP are cumulative with the effects seen in studies of DBP, DiBP BBP, and DEHP.

Assumption 2: Combination effects of phthalates with other anti-androgens can be approximated by using dose addition.

Mixtures assessments have been conducted primarily with those phthalates that influence male reproductive development to determine if those are additive in nature, specifically if they display dose addition at doses well above estimated human exposures. A single study has been conducted *in vivo* which tested the interaction effect of DINP and DEHP on testicular testosterone production (Borch *et al.*, 2004). Thirty-two dams were dosed with either 300 mg DEHP/kg bodyweight per day, 750 mg DINP/kg bodyweight per day, or a combination of these doses. Male fetuses were examined on gestation day 21, and blood and testes were collected for hormone analysis. The authors reported that a factorial statistical analysis revealed no statistically significant interaction between the effects of DEHP and DINP. In contrast, the assumption of dose-addition appears to be supported by the mixtures studies with phthalates such as DEHP and DBP, again using doses at or near the observable effect region (Ghisari and Bonefeld-Jorgensen, 2009; Howdeshell *et al.*, 2007; Howdeshell *et al.*, 2008a; Howdeshell *et al.*, 2008b; Jarfelt *et al.*, 2005; Martino-Andrade *et al.*, 2008; Rider *et al.*, 2008; Rider *et al.*, 2009).

The assumption of dose addition as the basis for conducting a cumulative risk assessment for humans is highly conservative (*i.e.*, dose-addition is assumed at levels below a threshold of response) and not well supported in the published literature. As stated by Borgert *et al.* (2004), dose addition may be a conservative assumption [for some effects] of chemicals when they are present at concentrations at or above their NOAELs, but that independence becomes more predictive when the concentrations of the component chemicals are well below their individual NOAELs. It is important to point out that the reason that components of mixtures may be less than additive when tests are conducted at low levels is that the modes of action could be different at different exposure levels. In particular, substances are much more likely to cause toxicological effects at exposure levels that overwhelm clearance mechanisms.

Borgert *et al.* (2004) also indicate that it is premature to assume dose addition for chemicals that appear to be mechanistically similar and to assume response addition models only for chemicals that appear to be mechanistically dissimilar. Because these simple models were developed for binary mixtures, their applicability to more complex mixtures is uncertain. Dose addition should be correlated with specific mechanistic features for particular toxic effects before the approach is generalized.

1.4.2 Conservative assumptions in the HI should be acknowledged.

(i) Dose-addition (DA) – Dose addition is based on the idea that all components in a mixture behave as if they are simple dilutions of one another. DA implies that every toxicant in the mixture contributes, in proportion to its toxic unit, to the overall mixture toxicity. This oversimplification introduces a high degree of conservatism and uncertainty.

(ii) Use of No Observed Effect Level (NOAEL)/Lowest Observed Effect Level (LOAEL) to describe dose-response data – Point estimates, such as NOAELs and LOAELs, neither represent effect concentrations or effect levels. Both are empirically based on experimental design and may not be accurate representations of the intrinsic hazard value of a chemical. Since point estimates do not represent equi-effective doses, the use of them in a CRA introduces an additional layer of conservatism and uncertainty into the HI approach.

(iii) Modified Points of Departure (MPOD) – Adjustment/uncertainty factors used in the calculation of the MPOD are quantitative judgments of qualitative deficiencies in the database and are typically based on default values. The use of these uncertainty factors results in the conservative estimate of an MPOD, and by extension, a conservative HI value.

1.5 Interpretation of Cumulative Risk Assessment

1.5.1 DEHP is the major contributor to the HI.

Because DEHP makes such a large contribution to the CRA, any reduction in exposure of any other phthalate would have a minimal impact on the overall hazard index. For example, for pregnant women in NHANES 2005-06 data (CHAP Table 2.16 and Table S-1), the estimated 95th percentiles for HQs of DINP are 0.01, 0.1 and 0.02 respectively for case 1, case 2 and case 3. However, the estimated 95th percentiles for HQs of DEHP are 6.0, 3.6 and 3.6 respectively for case 1, case 2 and case 3, *i.e.*, almost the same as the estimated 95th percentiles for HIs for case 1 (6.1), case 2 (3.7) and case 3(3.6). The similar pattern is observed for other percentiles of HIs and HQs.

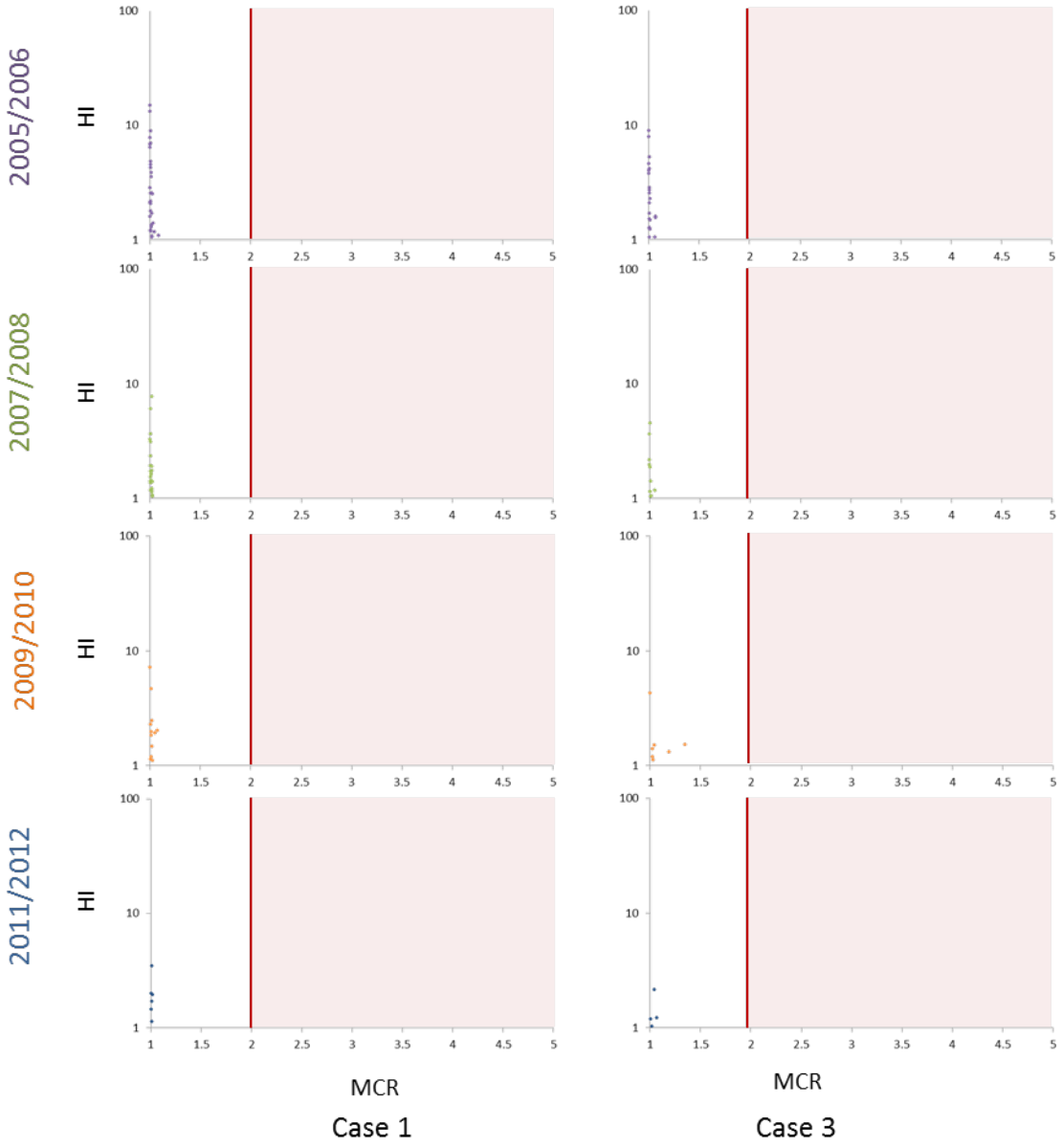
When assessing the results of a cumulative risk assessment it is important to consider not only whether the Hazard Index is above one, but also the general contribution of each substance to the overall risk. This is because in many cases the outcome of a cumulative risk assessment is driven by a single chemical (Price *et al.* 2011). As shown below, using the methodology of Price *et al.* (2011), it is apparent that DEHP is by far the principal contributor to the CRA. Accordingly, the only way to substantially impact the CRA is to continue to seek opportunities to reduce exposure to DEHP.

As defined by Price *et al.* (2011):

- Maximum cumulative ratio (MCR) is a method to determine relative contribution (Price *et al.* 2011)
    - MCR = HI/Max HQ
- As MCR values approach 1, the cumulative risk assessment (CRA) is being driven by a single chemical
- A threshold MCR value of 2 has been proposed (Price *et al.* 2012a, Price *et al.* 2012b)
- Risk management for all components of CRA approach needs to be considered I HI > 1 and MCR > 2
    - MCR < 2 indicates a single substances is responsible for 50-100% of the risk

In this particular case, given five phthalates were included in the CRA, an MCR = 5 would indicate equal contribution of all phthalates. An MCR = 1 would indicate a risk contributed by a single phthalate. Risk management measures on a cumulative basis are suggested to be considered when and MCR is greater than 2 (50% - 20% by a single phthalate in this case) (Price *et al.* 2012b). In Figure 6 below each point is the coordinate of a single individual from the population "women of reproductive age" positioned by here her HI and MCR value. Only individuals with a HI greater than 1 are pictured. As can be seen in the figure below for Case 1 and Case 3 each individual with an HI above 1 has an MCR value approaching 1. This result means nearly all the risk is due to a single phthalate. No individuals fall into the region that gives concern for combined effects (pink zone).

Figure 6 CRA for five phthalates indicates risk is due to single phthalate. No values HI>1 and MCR>2 (pink shaded region)



Data points represent women of reproductive age (15-45) from the NHANES datasets indicated (e.g., NHANES dataset from 2011/2012).

1.5.2 DINP is only analyzed for a small subset of the SFF populations

It is of interest to note that a limited subset of infants reported on in the CHAP from the SFF study include DINP metabolite data. This difference in sample number between DINP and the other phthalates used in the CRA is inconsistently noted. For example Table 2.5 of the CHAP report notes concentrations of DEHP and DINP metabolites indicating a sample size of 291 infants for the CHAP/SFF study, however, in Table 2.7 when reporting daily intake levels the report makes a subnote for DINP that the sample size is only 67 infants. More importantly when

reporting on the HI calculations for the SFF infants the report fails to note how the DINP hazard quotient (HQ) was incorporated. Appendix D indicates a sample size of 251 for HI calculations for the infant SFF population. Given that it is not possible for all those values to contain a HQ for DINP it is unclear how these analyses can be used for a recommendation for DINP.

1.6 The Analysis and Conclusions are Based on Outdated Exposure Information

As stated by the CHAP (CHAP, 2014, page 41, appendix D) "A limitation of the analyses presented here is the use of exposure data from 2005–06 for NHANES and 1999–2005 for the SFF. Since these data were collected, the Consumer Product Safety Improvement Act restricted some of the uses of the five phthalates evaluated. The impact on exposure is unknown and not accounted for in the calculation of the HI."

Data from 2011-12 for NHANES show that urinary metabolites of DEHP are approximately a factor of 3 lower than in the 2005/2006 data set used by CHAP in their cumulative risk assessment. Information summarized in Table 1 (below) documents that there has been a decrease in exposure to DEHP. As described elsewhere in this document, the use of the 2005/2006 survey information as opposed to more recent data has a significant impact on the calculation of the HI. Because DEHP is the major contributor to the HI, the use of outdated exposure information results in HI values that overestimate the cumulative risk.

Table 1. Geometric means (95% confidence interval) of urine concentrations (μg/L) for the U.S. population from the National Health and Nutrition Examination Survey (NHANES).

Survey years	Mono-(2-ethyl-5-carboxypentyl) phthalate	Mono-(2-ethyl-5-hydroxyhexyl) phthalate	Mono-(2-ethyl-5-oxohexyl) phthalate	Mono-2-ethylhexyl phthalate
99-00				3.43 (3.19-3.69)
01-02		20.0 (17.8-22.5)	13.5 (12.0-15.0)	4.27 (3.80-4.79)
03-04	34.7 (31.0-38.9)	21.7 (19.3-24.4)	14.5 (13.0-16.1)	2.34 (2.10-2.62)
05-06	38.6 (34.7-42.9)	25.5 (23.0-28.2)	16.2 (14.6-18.0)	3.04 (2.78-3.32)
07-08	33.3 (28.7-38.6)	22.1 (18.7-26.0)	12.2 (10.3-14.3)	2.64 (2.29-3.05)
09-10	20.7 (18.5-23.3)	12.9 (11.3-14.7)	8.02 (7.11-9.06)	1.59 (1.41-1.79)
11-12	12.9(12.0-13.9)	7.90(7.47-8.35)	5.08(4.77-5.41)	1.37(1.25-1.49)

1.7 The use of the CHAP CRA Differs from Normal Practice as Outlined by the World Health Organization (WHO)

The WHO Risk Assessment Framework for combined exposure to multiple chemicals was designed to aid risk assessors in identifying priorities for risk management where co-exposures are expected.

Below is discussion of the WHO framework and how the CHAP CRA fits into that framework.

(a) Purpose and Focus of a Framework Assessment

- problem formulation (PF) requires preliminary consideration of hazard characterization and exposure assessment as a basis to plan the risk assessment process.

  Consideration was not given by the CHAP specifically to hazard characterization. Problem formulation was not a consideration although it should have been, given that the CHAP was only instructed to "consider" cumulative effects. Problem formulation could have achieved that aim.

- key considerations for PF would include: What is the nature of exposure? Are the key components known? Are there data available on the hazard of the mixture itself? WHO framework indicates that lack of this information on these aspects precludes a framework analysis.

  As described in more detail elsewhere, the CHAP had data to support the nature of the exposure (*i.e.*, continuous, low dose, co-exposure). The components are known to co-occur based on human biomonitoring data. Data on the mixture itself does not exist. Data on various phthalates mixtures have been created and tested at concentrations at or just below the NOELs. Mixture effects observed can be described equally well as response addition and concentration addition (Borgert *et al.*, 2012). Mixture effects at doses well below the NOEL are unknown. Potentially, mode of action (MOA) is different at high and low doses. Therefore, even though some data exist on phthalate mixtures at or near effect levels, caution should be taken in extrapolating the results to human-relevant doses.

- Key considerations include: Is exposure likely, taking into account the context? Is there a likelihood of co-exposure within a relevant timeframe?

  Given biomonitoring data the CHAP was able to conclude exposure was likely, given the context, and co-exposure was likely within a relevant timeframe

- Key considerations included: What is the rationale for considering compounds in an assessment group?

    Normally, this is based on predictive information on chemical structure, and or hazard data (e.g., similar MOA, effects observed in the same target organ, same biological outcome?) The CHAP relied on hazard data suggesting that the phthalates included in the assessment group all result in some form of "phthalate syndrome" effects. As described in more detail elsewhere in this document, data from Clewell et al (2014) demonstrate that DINP does not induce the “phthalate syndrome”. Therefore, the scientific basis for a common assessment group is questionable.

(b) <u>The WHO Framework</u>

- The initial tier begins with simple but conservative assumptions for both exposure and hazard. These assumptions are refined and replaced with increasingly detailed data and models, but only if there is an indication of concern or excessive risk.

- In the Tier 0 Hazard assessment, conservative early assumptions for an assessment group should be considered together such as due to similar target organ, similar mode of action, based on predictive hazard tools, and in the absence of information on individual components. All components are assumed to have the same potency as the most toxic compound.

- In the Tier 1 Hazard assessment, the analysis is refined by incorporating additional information on the potency of individual components of the common effect and more accurate measures of points of departure for hazard.

    The CHAP based their assessment group on a similar group of effects (phthalate syndrome). The CHAP did not have MOA data or a single endpoint to justify the common assessment group. The CHAP assumed 4 of the 5 phthalates included in the assessment group were equipotent to DEHP in case 2. Only the potency of DINP was estimated to be less than that of DEHP (based on a single study by Hannas *et al.*, 2011). Therefore, based on the WHO framework, the CHAP conducted a Tier 0/1 hazard assessment (*i.e.*, no single endpoint, but common target system (dev. system) (tier 0) and some consideration of potency (tier 1))

- Tier 3 Exposure assessment: estimates of exposure are probabilistic in nature, taking into account distributions of exposure factors or exposure data. This approach requires

representative information on exposure for the scenario of interest for the relevant populations for different uses and across populations. Models often include multiple-source exposures.

As the CHAP had urinary metabolite data available for the phthalates in the common assessment group, a multiple-source exposure estimate could be obtained for the populations of interest (*i.e.*, women of reproductive age) and essentially a Tier 3 assessment was conducted**.** Infants were also included although they were not specifically identified as a sub-population of concern.

- The framework indicates that for risk characterization, the nature of considerations that constituted the basis for determining that a higher-tier assessment is required is explicitly stated (*i.e.*, adequacy of the margin of exposure in the context of uncertainty associated with both estimated exposure and hazard).

  The CHAP did not adequately describe the nature of the uncertainty with their approach and did not contextualize the conservatism in a Tier 0/1, Tier 3 cumulative assessment. Conservatism inherent in this approach has been described previously by Borgert et al (2012).

Overall, the cumulative RA conducted by the CHAP can be considered a screening level assessment. As problem formulation was not conducted, the goal of conducting the assessment was never specifically defined (*i.e.*, who is being protected? at what level? over what timeframe?) hence making conclusions from a screening level and highly conservative assessment untenable. Instead additional iterations with higher tier assessments especially for hazard should have been undertaken before recommendations were made.

2. Issues Related to Data Quality

2.1 Rigor of the Review of the Literature Review

A systematic review of the literature includes evaluation of the quality, reliability, and appropriateness of each study as well as inclusion and exclusion criteria for assessing relevance of a study to the evaluation of risk. Clear procedures and protocols should be provided articulating the basis for these determinations to ensure transparent, consistent and scientifically sound evaluation of the data. Without such an evaluation, studies of lower quality are inappropriately accorded too much weight in the overall assessment, leading to a flawed evaluation. It is critical that CPSC rely on the studies that are of the highest quality, not simply those studies that produce the lowest points of departure.

The CHAP addressed study selection in two notable places in the document. In the first mention (p5), the CHAP addressed study inclusion/exclusion as follows *"[i]n cases in which peer-reviewed data were not available, the CHAP made decisions on a case-by-case basis as to whether non–peer-reviewed data would be used in making recommendations to the CPSC"*. As no additional information is provided regarding the criteria that were considered as the basis for data selection, the reader is left to assume a subjective approach to study selection was applied. In the second mention of study selection criteria (section 2.3.2), the CHAP lists a series of questions used for key study selection. These questions focused solely on study design (e.g., number of dose levels, number of animals) and did not include consideration of model relevance or database consistency. These questions do not address the adequacy of data in a systematic manner. Furthermore, there is no transparency with respect to the answers to these questions for each individual study. Therefore, it is unclear if all studies were evaluated systematically and equally or if the data were subjectively selected in support of their conclusion.

Incorporating a clearly defined systematic approach into influential scientific assessments is consistent with the recommendations made by the National Academy of Sciences to the USEPA in their 2011 review of the US EPAS draft IRIS assessment on formaldehyde.[18] In this report the committee made five general suggestions for improving the EPA's IRIS assessments one of which was to present clear and expanded descriptions of the rationale by which the studies upon which toxicity criteria are based are selected. This suggestion is currently being adopted into the USEPA's IRIS process[19], is included in the USEPA's risk framework[20], is relied upon by the NTP for drawing conclusions about potential human health hazards,[21] and should be followed for all regulatory agencies responsible for developing influential scientific assessments.

2.2 Lack of Weight of Evidence and Inadequate Integration of Data.

The CHAP failed to apply a scientifically-solid framework for integrating study results based on a weight of evidence approach for evaluating potential risks to humans at environmentally relevant exposures. Furthermore, they failed to apply a transparent and consistent approach for evaluating and integrating evidence using uniform evaluation methods to determine quality and reliability for the different types of studies, and addressing uncertainties.

---

[18] National Research Council. 2011. Review of the Environmental Protection Agency's Draft IRIS Assessment of Formaldehyde. Committee to Review EPA"s Draft IRIS Assessment of Formaldehyde. Board of Environmental Studies and Toxicology. Division of Earth and Life Sciences. Available at http://www.nap.edu/catalog.php?record_id=13142).

[19] National Research Council. 2014 *Review of EPA's Integrated Risk Information System (IRIS) Process*. Washington, DC: The National Academies Press.

[20] US EPA. 2014. Framework for Human Health Risk Assessment to Inform Decision Making. EPA/100/R-14/001

[21] Birnbaum LS *et al.* 2013 Implementing systematic review at the National Toxicology Program: status and next steps Environ Health Perspect. 2013 Apr;121(4):A108-9.

Failure to rely upon a weight of evidence approach is exemplified in the following three examples. First, and as discussed above, the articulated criteria for key study selection in section 2.3.2 focused solely on study design (e.g., # of dose levels, # of animals) and did not include consideration of model relevance or database consistency. Second, the criteria for study inclusion/exclusion were not transparently articulated as reflected in the following quote (p5) "*In cases in which peer-reviewed data were not available, the CHAP made decisions on a case-by-case basis as to whether non–peer-reviewed data would be used in making recommendations to the CPSC"*. And, third, database inconsistencies and strength of database were ignored when drawing conclusions. This is particularly evident when assessing and integrating the findings from epidemiological studies. In section 2.4, the CHAP acknowledges (section 2.4.1 p28) the weakness of the epidemiological database in at least two places with the following statements: "*the results of these [three cohort studies] were not entirely consistent*" and "*.. the data on phthalates and AGD are suggestive and human data suggest that AGD is a relevant marker for reproductive health outcome*". However, despite this acknowledgement the CHAP concluded the following "*based on the human data on gestational exposure and reduced AGD, exposure to DEP, DBP and DEHP metabolites should be reduced*", therein ignoring the weaknesses they had previously identified.

In accordance with established best practices of systematic evidence-based reviews, the CHAP should employ a consistent weight of evidence framework, based on specific hypothesized Modes of Action (MOAs) to permit data from laboratory experiments, epidemiological investigations, and mechanistic research to be integrated in a manner that provides a robust understanding of the potential hazards and risks that exposures to a substance could pose to humans. Embedded in this framework should be a clear and systemic approach for addressing the uncertainties of the data equally.

3. Issues Related to the Selection of "Anti-Androgenic Effects" as the Key Toxicological Endpoints for Cumulative Risk Assessment

3.1 Selection of Endocrine-Mediated Effects as the Key Toxicity Endpoints for Cumulative Risk Assessment

The CHAP selected endocrine-mediated effects as the key endpoints for the cumulative risk assessments. However, their rationale for this choice, as stated on pg 13 of the CHAP report is misleading: *"The most sensitive and most extensively studied endpoint is male developmental toxicity in the rat, and therefore the CHAP focused on this toxicity endpoint, consistent with the stance taken in earlier assessments by other bodies (National Research Council [NRC, 2008])."* In fact developmental toxicity is not the most extensively studied endpoint or always the most sensitive. Male developmental toxicity in the rat is the most sensitive endpoint for some phthalates such as DEHP, but is not the most sensitive endpoint for all phthalates. More specifically, male reproductive toxicity is not the most sensitive endpoint for DINP. This was pointed out in the report on pg 98… *"The NOAEL for liver toxicity for DINP (12 mg/kg-day), as for DIDP, is lower than the lowest NOAEL for anti-androgenic toxicity (50 mg/kg-day for MNGs)…"* Parenthetically it should be noted there is evidence that the multinucleated gonocytes

(MNGs) are not a consequence of reduced testosterone synthesis. For example mice lack the associated anti-androgenic effects demonstrated in rats, yet produce MNGs after phthalate exposure. Additionally, an investigation into the role of androgens in fetal testis development and dysgenesis concluded that the induction of MNGs was mechanistically separated from intra-testicular testosterone reduction (Scott *et al.* 2007). Finally, of importance is that MNGs are not considered adverse as they are eliminated in a p53-dependent manner from the seminiferous epithelium within 1–2 weeks postnatally (Johnson *et al.* 2012).

4. The Data do not Justify the Inclusion of DINP in the Cumulative Risk Assessment

The information provided below demonstrates that the inclusion of DINP in a cumulative risk assessment for the adverse outcomes from the anti-androgenic effects of phthalates is not appropriate. In addition given the evidence that humans are at least less sensitive (if not non-responsive), and potentially refractory to the effects seen in rats, it calls into question the 100-500 fold uncertainty factors used for the points of departure used in the CHAP CRA. In essence the adjustment of the data in this manner assumes humans are potentially 100-500 times more sensitive to the effects seen in the animal studies. The data clearly support that this is not the case. Though the CHAP report authors assume relevance of the effects to humans, a refinement of the interspecies assessment factor from a default value of 10 (humans assumed more sensitive) to a more chemical specific assessment factor (CSAF) less than or equal to 1, would be appropriate for these substances. As shown below, an interspecies assessment factor of less than one is justified. In short, just based on species-differences, the cumulative risk is over estimated by a factor of at least 10 when applied to human risk. This conservative application of assessment factors, along with the use of outdated exposure information and incorrect points of departure, highlights that the recommendations given in the CHAP report are not supported by the data.

4.1 Rat Phthalate Syndrome

In the section on Rat Phthalate Syndrome the report states “Active” phthalates start with diisobutyl phthalate (DIBP, with three carbon atoms in the alkyl backbone) and end with DINP (with~seven or eight carbons in the alky chain backbone).” (pg. 16 CHAP report)

The term “rat phthalate syndrome” was coined to encompass a group of effects observed in male rats from exposures during the critical window of male reproductive tract development (Gray and Foster, 2003). However, the basis for classifying this group of effects as a “syndrome” specifically attributable to phthalates as a class is weak and imprecise. There are significant differences in toxicity between the low molecular weight phthalates (LMW) and the high

molecular weight phthalates (HMW), such as DINP and DIDP.[22] When these differences in toxicity are appropriately taken into consideration, it is clear that the inclusion of DINP in a cumulative risk assessment based on the "rat phthalate syndrome" is not warranted, since there is no induction of the adverse outcomes of maldevelopment of the male reproductive tract that are observed with certain other phthalates.

4.1.1 "Rat Phthalate Syndrome" – A Hypothesis for LMW Phthalate-Induced Male Reproductive Tract Effects

The group of effects induced by some phthalates, which has led to a hypothesis that these effects are due to a common, endocrine-mediated process, has collectively been described by some researchers as the "rat phthalate syndrome". These effects, as defined by Gray and Foster (2003) include: decreased anogenital distance, nipple retention, infertility, decreased sperm count, cryptorchidism, hypospadias, and other reproductive tract malformations such as testicular, epididymal, and gubernacular cord agenesis. The validity of this hypothesized syndrome for use in a phthalate cumulative risk assessment is questionable.[23] A control incidence of this syndrome has never been established and the threshold for inclusion based on incidence and severity of each effect has never been defined, though it has been suggested that one effect or merely a proposed sentinel event is enough to warrant inclusion. In addition, a number of non-phthalate compounds induce one or more of the included effects which belies the specificity of this description to "phthalates" only.

Furthermore, while these effects are observed with some phthalates, a weight of the evidence review of all available data indicates that DINP does not induce the effects characteristic of the "rat phthalate syndrome". The phthalates DiBP, DBP, BBP, and DEHP clearly induce the effects characterized as the "rat phthalate syndrome": hypospadias, cryptorchidism, decreased anogenital distance, nipple retention, changes in androgen sensitive tissue weight and infertility. As demonstrated below, DINP does not cause these effects. Thus, it is inappropriate to name this group of effects as a syndrome attributable to phthalates as a class.

Moreover, the mode(s) of action leading to the observed effects included in the hypothesized "rat phthalate syndrome" is not known. A molecular target(s) of the phthalates has not been identified and likely differs based on the phthalate (*i.e.*, pharmacodynamic differences). A reduction of fetal testosterone and/or a reduction in insulin-like 3 peptide hormone biosynthesis (insl3) during the critical window of male reproductive tract development have been

[22] Low molecular weight (LMW) phthalates are those with 4-6 backbone carbons in their alkyl side chains, e.g., DBP, BBP, DEHP. High molecular weight phthalates (HMW) are those with > 6 backbone carbons, e.g., DINP, DIDP.

[23] For this reason we put quote marks around the term "rat phthalate syndrome". Another term that has been proposed is androgen insufficiency syndrome, which has the merit of not being overbroad with respect to phthalates and underbroad with respect to other chemicals. However, as discussed below, each effect is not necessarily related to androgen levels.

hypothesized to be critical contributors or common key events predictive of the "rat phthalate syndrome"; each is discussed in more detail below. However, a number of non-overlapping disrupted pathways may result in the varied and complex responses. This complexity highlights the need to carefully examine the specific toxicity, adverse health effects, and associated events for each individual phthalate. Of note is the appropriateness of grouping endpoints for the cumulative risk assessment. The complexity of "rat phthalate syndrome" makes the appropriate selection of the endpoint upon which to base a cumulative risk assessment difficult and the selection of earlier "biomarker" events (case 3) a highly conservative, potentially speculative, assessment.

Given the differences noted between phthalates, it is plausible that multiple modes of action may be at play; observation of a single precursor event (e.g., reduced testosterone) may not be predictive of the suite of effects described above, as exemplified by the data available for DINP.

4.1.2 Role of insl3 in "Rat Phthalate Syndrome"

Insulin-like hormone 3 (insl3) is a peptide hormone produced by the Leydig cells of the testes which has been shown to be associated with gubernacular defects and cryptorchidism when reduced (Adham *et al.*, 2000; Nef and Parada, 1999; Zimmermann *et al.*, 1999). Specifically, insl3 induces the gubernacular cord to differentiate and mature, thus facilitating the first phase of testes descent from the kidney area to the inguinal region during fetal life (Zimmermann *et al.*, 1999). Mice without a functional insl3 gene display cryptorchid testes and normal androgen levels. Androgen also plays a role in testis descent by acting to regress the cranial suspensory ligament during the first phase of testis descent. In the untreated (control) female rodent fetus, the gubernacular cord involutes in the absence of insl3 and the cranial suspensory ligament develops in the absence of testosterone to maintain the position of the ovaries near the kidneys (Howdeshell *et al.*, 2008a).

Two studies have examined the effect of DINP on insl3 mRNA levels. In one study, an increase in insl3 mRNA was observed 2 days following the last dose (*i.e.*, GD 19.5) of DINP. However, the authors suggested that the increase may have been due to a "rebound effect" from the low testosterone production at the time dosing was initiated (*i.e.*, GD 13.5) (Adamsson *et al.*, 2009). Results of a second study were presented in a poster recently at the 2011 Society of Toxicology meeting; preliminary data suggested DINP did not affect insl3 mRNA levels (Lambright *et al.*, 2011). Therefore, along with evidence from the definitive 2-generation study and developmental toxicity studies where cryptorchidism was not observed (see below), DINP likely does not affect insl3.

4.1.3 Role of Fetal Testosterone in "Rat Phthalate Syndrome"

In order to assess the role of altered fetal testosterone as a critical contributor or common key event predictive of the "rat phthalate syndrome", current knowledge of the role of testosterone in the developing male fetus needs to be understood. Steroidogenesis in the fetal rodent and human testis has been reviewed in detail (Scott *et al.*, 2009), and key events are described here.

4.1.3.1 Altered testosterone levels in the rat fetus may be due to growth and differentiation factors (paracrine factors)

Beginning at gestational day (GD) 14.5 to 15.5, testicular testosterone production is initiated in the rat (Habert and Picon, 1984; Warren *et al.*, 1972). The mechanism for initiation is somewhat unclear as luteinizing hormone (LH) secretion, a primary stimulatory hormone, does not start until embryonic day 17.5 (Aubert *et al.*, 1985). This suggests that testosterone production is largely regulated either autonomously or by paracrine factors during embryonic days 15.5 – 17.5 (Scott *et al.*, 2009). This time period has been termed the "masculinization programming window" and is thought to be the critical window for androgen influence necessary for morphological differentiation of the male genitalia (e.g. epididymis, vas deferens, seminal vesicles, prostate, penis, scrotum and perineum) (Scott *et al.*, 2009). Following this programming window and a peak in fetal testosterone on approximately embryonic day 18 (Livera *et al.*, 2006), LH levels begin to rise and influence gonadotropic function. Based on these events and given the most common dosing regimen (*i.e.* single or repeated dose during GD 7 - GD 21) in short term *in vivo* rat studies, altered testosterone levels may be a result of disrupted paracrine factor action and or influence (Scott *et al.*, 2009).

4.1.3.2 Humans differ from rats in aspects of testicular steroidogenesis

Fundamental control of steroidogenesis in the fetal rat differs from that in the human fetus. This point is important since it is frequently claimed that the pathway (sexual differentiation) that phthalates disrupt in the fetal male rat is highly conserved in all mammals and is known to be critical for human reproductive development. Indeed, commonalities exist between humans and rodents during the period of sexual differentiation (*i.e.* the time when a fetus can be morphologically distinguished as being male) and to some extent masculinization. However, a clear difference is noted in the stimulatory mechanisms for testicular steroidogenesis during the critical period when masculinization of the reproductive tract is being programmed. As described for the rat, the 2 day time period (GD 15.5-17.5) during which testosterone is produced and masculinization occurs is largely LH-independent (Scott *et al.*, 2009). Human fetal testosterone production begins around gestational week 8 and is mainly controlled by chorionic gonadotropin (hCG), a hormone not produced by rodents. By gestation week 12, hCG begins to decline and LH levels are seen to rise, although hCG is two to six times more potent than LH on a weight basis and may continue to strongly stimulate steroidogenesis through week 20 (Dufau *et al.*,

1972; Lee and Ryan, 1973). Unlike rodents, paracrine factors likely have a secondary or supporting role in human testosterone secretion and do not initiate production.

Basic differences in the steroidogenic cascade are also noted. The principle form of circulating cholesterol differs between rats and humans. HDL is the primary source taken up by the SRB-1/HDL receptor on the Leydig cell in rats and LDL is the primary source taken up by the LDL receptor on the Leydig cell in humans. In addition, the preferred steroid biosynthetic pathway converting cholesterol to testosterone differs; the Δ4 pathway (*i.e.* progesterone and its intermediate 17α-hydroxyprogesterone) predominates in rats while the Δ5 pathway (*i.e.* pregnenolone and its intermediates, 17α-hydroxypregnenolone and DHEA) is the predominant mechanism of testosterone synthesis in humans. These differences must be considered when characterizing the relevance of reported rodent effects and their extrapolation to human hazard characterization and risk assessment.

4.1.3.3 Existing data do not support relevance to humans of reduced fetal testosterone in rats

Species differences in response to phthalates have become more apparent in the recent literature. *In utero* exposure of mice and rats to DBP results in multinucleated germ cell formation and an increase in seminiferous tubule diameter, yet rats only exhibit suppression of fetal Leydig cell steroidogenesis (Gaido *et al.*, 2007). This difference could be a species specific effect of DBP exposure on fetal Leydig cell SREBP2 activity; however the underlying mechanism is unknown (Johnson *et al.*, 2011).

Limited data have been reported from studies in which effects of phthalates have been tested on human fetal testes. Lambrot *et al.*, 2008 investigated the effect of MEHP on human fetal testes recovered during the first trimester (7-12 weeks) of gestation. MEHP had no effect on basal or LH-stimulated testosterone and did not affect proliferation and apoptosis of Sertoli cells. Reduced mRNA expression of anti-Müllerian hormone was reported and a reduced number of germ cells (via increased apoptosis) were also seen. Similarly, Hallmark *et al.* (2007) reported no effect on human fetal testis explants cultured with 10^{-3}M MBP for up to 48hrs. This included measurement of intra-testicular testosterone levels and cytochrome P450 side chain cleavage enzyme expression as well as Leydig cell aggregation. However, the authors of the paper questioned the utility and validity of the *in vitro* system. Human fetal testes have also been xeno-transplanted within the renal subcapsular space of a nude rat host followed by three days exposure to DBP (Heger *et al.*, 2010, 2011). Results, which were presented to the CHAP, indicate DBP did not affect steroidogenic gene expression. An increase in multinucleated gonocytes (MNGs) per total number of germ cells was reported although the significance of this effect is not known. Therefore, limited data using human tissue has not indicated any effect by phthalates on the Leydig cell or suppression of testosterone. This highlights the need for further

research but also calls into question the relevance of testosterone reduction in rats by phthalates for human health risk assessment.

4.2 DINP Does Not Induce "Rat Phthalate Syndrome"

The following section reviews the available data from studies which have specifically investigated DINP, including data on a suggested critical contributor, testosterone reduction, and each of the effects proposed to be within the hypothesized "rat phthalate syndrome". Infertility, the most severe outcome of disruption of male reproductive tract development, is also discussed.

In addition, discussed in more detail below, robust developmental studies of DINP, consisting of a gavage study using 144 pregnant rats and a dietary study using 100 pregnant rats, were recently published by the Hamner Institute (Clewell *et al.*, 2012a, b). These studies were designed to provide strong statistical power for analyzing, collectively, the kinetics and fetal testes effects of DINP and post-natal effects including nipple retention and AGD as well as any malformations of the male reproductive tract including hypospadias, cryptorchidism, and epididymal malformations, both gross and histological and the endpoints attributed to the hypothesized "rat phthalate syndrome." Investigation of effects at GD 19 gave a no observed effect level (NOEL) of 50 mg/kg/day based on increased MNGs and reduced testes testosterone concentration in the fetal rat. There is evidence that the multinucleated gonocytes (MNGs) are not a consequence of reduced testosterone synthesis. For example mice lack the compliment of antiandrogenic effects demonstrated in rats, yet produce MNGs after phthalate exposure. Additionally, an investigation into the role of androgens in fetal testis development and dysgenesis concluded that the induction of MNG's was mechanistically separated from intra-testicular testosterone reduction (Scott *et al.* 2007). Finally, of importance is that MNGs are not considered adverse as they are eliminated in a p53-dependent manner from the seminiferous epithelium within 1–2 weeks postnatally (Johnson *et al.* 2012). As noted by the authors, a No Observed Effect Level (NOEL) of 50 mg/kg/d was determined in Clewell *et al.* given no adverse effects were seen at any dose tested up to approximately 750 mg/kg/d. DINP has not been shown to induce permanent alterations in the male reproductive tract or fertility at doses that are well in excess of 50 mg/kg/day. Additionally, in these studies global endpoint analysis showed no evidence of a rat "phthalate syndrome" on PND 49 with DINP administration. (Clewell *et al.* 2012a; Clewell *et al.* 2012b)

4.2.1 DINP Induces a Transient Decrease in Fetal Testosterone Levels in High Dose Gavage Studies

Several short term *in vivo* studies have been conducted in rats that specifically evaluated the potential for DINP-induced effects on plasma/testicular testosterone production or content (Adamsson *et al.*, 2009; Boberg *et al.*, 2011; Borch *et al.*, 2004; Gray *et al.*, 2000; Lee *et al.*, 2006a; Lee *et al.*, 2006b; Clewell *et al.*, 2012a; Clewell *et al.*, 2012b). For comparison, the

results of those studies are summarized (Table 2). Of those, two studies, one examined only a single dose of DINP, and in the other, effects were observed only in one dose group in the middle of the dose range (Boberg *et al.*, 2011; Borch *et al.*, 2004). Four studies reported no effects for various testosterone measurements at multiple time points following exposure (Adamsson *et al.*, 2009; Boberg *et al.*, 2011; Gray *et al.*, 2000; Lee *et al.*, 2006a; Lee *et al.*, 2006b). Of the remaining two studies one indicated a dose response decreases 2 hours after dosing ceased, however by 24 hours post dosing the testosterone levels had rebounded and were higher than controls (Clewell *et al.*, 2012a). The final study measured testosterone levels in adulthood, following early life exposure (gestation thru weaning). This study demonstrated no differences in testosterone levels, indicating any potential reduction after *in utero* exposure is transitory and does not persist later in life (Clewell *et al.*, 2012b).

While two studies have reported an effect on fetal testosterone levels at GD 21, limitations of the studies should be taken into consideration. Both studies that reported an effect used high doses of DINP (e.g. 750 mg/kg/day). In addition, a clear dose-response was not demonstrated (Boberg *et al.*, 2011). At times after GD 21, no effects on fetal testosterone levels were observed, indicating the reductions observed at the early time point were transient.

Table 2 - - Studies that examined DINP effects on plasma/testicular testosterone production or content

					Testosterone Concentrations			
	Route/Strain	Dose (mg/kg)	Exposure Duration	Testosterone Measurement	Blood Serum	Blood Plasma	Intratesticular content	Testicular production
(Gray *et al.*, 2000)	G/SD	750	GD 14 – PND 3	PND 90	**No effect**	n.d.	n.d.	n.d.
(Borch *et al.*, 2004)	G/W	750	GD 7 – GD 21	GD 21	n.d.	**No effect**	**(+) approximately 60% reduction**	**(+) approximately 60% reduction**
(Lee *et al.*, 2006a)	D/W	5 50 500 1100	GD 15 – PND 21	PND 140	**No effect**	n.d.	n.d.	n.d.
(Lee *et al.*, 2006b)	D/W	5 50 500 1100	GD 15 – PND 21	PND 7	**No effect**	n.d.	n.d.	n.d.
(Adamsson *et al.*, 2009)	G/SD	250 750	GD 13.5 – GD 17.5	GD 19.5	n.d.	n.d.	**No effect**	n.d.
(Boberg *et al.*, 2011)	G/W	300 600 750 900	GD 7 – PND 17	GD 21	n.d.	**No effect**	**(+)approximately 40% reduction (600 mg/kg only)**	**No effect**

(Boberg *et al.*, 2011)	G/W	300 600 750 900	GD 7 – PND 17	PND 90	n.d.	n.d.	**No effect**	n.d.
(Clewell *et al.*, 2012a)	G/SD	50 250 750	GD 12 – GD 19	GD 19	n.d.	n.d.	**(+)approximately 50% reduction 2h post treatment approximately 60% increase 24h post treatment (250 mg/kg)**	n.d.
(Clewell *et al.*, 2012b)	D/SD	50 250 750	GD 12 – PND 14	PND 49	n.d.	n.d.	**No effect**	n.d.

G: Gavage, D: Diet, SD: Sprague-Dawley, W: Wistar, n.d.: no data

4.2.2 DINP dDoes Not Induce Permanent Changes in Anogenital Distance

Anogenital distance (AGD) is a sexually dimorphic trait in laboratory rodents and humans; rodent males exhibit a distance 2 – 2.5 fold greater than females. Androgens are responsible for normal AGD elongation in neonatal males (Clemens *et al.*, 1978; Hotchkiss *et al.*, 2007; Imperato-McGinley *et al.*, 1985). In laboratory animals, agents that are androgen receptor antagonists will induce a decrease in AGD in males.[24]

Anogenital distance was reported to be unaltered in two studies in which: a single dose of 750 mg/kg/day DINP was administered by gavage (Gray *et al.*, 2000); doses up to ~2500 mg/kg/day were administered via the diet (Masutomi *et al.*, 2003).

Boberg *et al.* (2011) reported a small (6%) but statistically significant decrease in anogenital distance in males exposed to DINP at 900 mg/kg/day on post-natal day 13. However, the authors reported there was no difference between treated animals and controls on post-natal day 90 and suggested that the change in AGD was transitory.

Clewell *et al.* (2012b) reported no statistically significant decreases in anogenital distance in males exposed up to 750 mg/kg/day on post-natal day 2 and 49. There was a slight statistically significant difference at post-natal day 14 in the highest dose group. Anogenital difference is highly dependent on animal size. At post-natal day 14 the pup weights for males in this group were also statistically different than controls, however these animals were no longer different for either weight or anogenital distance at post natal day 49. Given that anogenital differences

[24] As described in Attachment A, DINP is not an androgen receptor antagonist (Takeuchi *et al.*, 2005).

induced by anti-androgenic influences *in utero* would already be apparent at birth, the difference at post-natal day 14 was likely due to a difference in pup size, and not evidence of an anti-androgenic effect. This conclusion is supported by the lack of difference at post-natal day 2 and further supported by the return to control values by post-natal day 49.

Lee *et al.* (2006b) reported a significant decrease in anogenital distance at all doses tested (0, 40, 400, 4000, or 20000 pm in the diet on GD 15 through PND 21) on post natal day 1. However, these results are suspect because of the very small difference between the control (2.5) and the treated (< 0.1 below 2.5) normalized values for all dose groups. This finding was reported as being statistically significant in each dose group, yet with a unit number potentially as low as 16 animals, the statistical findings seem suspect and draw into question whether this is a reporting error, especially since potent anti-androgens that were also studied in this report exhibited no effect for this measurement. As pointed out by Foster and McIntyre (2002), "a 2 to 3% change in anogenital distance although measurable is unlikely to be biologically of importance and in isolation would not necessarily be considered adverse".

Anogenital distance was specifically examined as part of the two-generation reproductive toxicity study protocol used for DIDP (Hushka *et al.*, 2001). DIDP (0.02, 0.06, 0.2 or 0.4% diet) did not affect AGD in the F_1 or F_2 pups when examined on post-natal day 0.

4.2.3 DINP Does Not Induce Permanent Nipple Retention

Nipple retention in males is thought to be a sensitive endpoint downstream of a reduction in fetal testosterone and has been assessed in several studies. As discussed earlier, further studies are warranted to determine if fetal reductions in testosterone are necessary and sufficient to produce this effect. The development of the rodent nipple is sexually dimorphic (Kratochwil, 1971; Kratochwil and Schwartz, 1976). Although mammary gland development begins similarly in both male and female rodent fetuses, offspring female rats and mice have nipples but males do not. In the developing rodent fetus, di-hydroxy testosterone produced locally from fetal testosterone causes regression of the nipple anlagen (Imperato-McGinley *et al.*, 1986; Kratochwil, 1977, 1986). This process can be disrupted, and these offspring subsequently display nipples. However, further studies are warranted to determine if fetal reductions in testosterone are necessary and sufficient to produce this effect.

As reported in Gray *et al.* (2000), data for DINP indicated that at 13 days of age, infant males with areolas were observed at an incidence of 22% compared with controls (0%). At approximately 5 months of age, 2/52 male pups displayed permanent nipples where the number of nipples equaled 1 and 6 for each of the two males. This effect was considered to be a malformation and was reported collectively with 2 other malformations as statistically significant, although the endpoint on its own was not statistically significant. The range of

historical control values is important for understanding the low incidence effects. In this study the control incidence for areola retention was reported to be zero, but in a subsequent study from the same lab using the same rat strain, control values are reported as 14% (Ostby *et al.*, 2001a) which confounds interpretation of the results of the earlier study.

Boberg *et al.* (2011) reported a significant increase in nipples in males exposed to DINP at 750 and 900 mg/kg/day (average of 3 nipples in each dose group) as compared to controls (average of 2 nipples) on post natal day 13. However, there was no difference in the number of nipples in males between control and treated animals on post natal day 90. Since nipple retention was not observed on post natal day 90, the utility of this endpoint for hazard assessment is questionable.

The biological and/or toxicological significance of nipple retention observed in early postnatal male rats is questionable. Studies examining the effects of in utero exposure to finasteride, a 5α – reductase inhibitor, demonstrated that finasteride exposure induced nipple/areola retention in perinatal male rats, but the effects were temporary (Clark *et al.*, 1990), similar to the finding of Boberg *et al.* (2011) and Carruthers and Foster (2005). Furthermore, unlike rats, human males do not lose their nipples, significantly challenging the relevance of this endpoint for use in human hazard assessment or by extension to cumulative risk assessment.

Clewell *et al.* (2012b) reported no significant difference in nipples in males exposed to DINP at approximately 50, 250, and 750 mg/kg/day. This study included 100 pregnant females and was designed to provide strong statistical power for analyzing post-natal effects including nipple retention.

4.2.4 DINP Does Not Induce Cryptorchidism, Hypospadias or General Reproductive Tract Malformations

Gross male reproductive tract malformations, such as cryptorchidism or hypospadias, have not been reported in any studies for DINP; including, the definitive two-generation reproductive toxicity studies (Hushka *et al.*, 2001; Waterman *et al.*, 2000), and a number of other *in vivo* studies previously mentioned (Adamsson *et al.*, 2009; Boberg *et al.*, 2011; Borch *et al.*, 2004; Gray *et al.*, 2000; Hellwig *et al.*, 1997; Kwack *et al.*, 2009; Lee and Koo, 2007; Lee *et al.*, 2006a; Lee *et al.*, 2006b; Masutomi *et al.*, 2004; Masutomi *et al.*, 2003; Waterman *et al.*, 1999).

Gray *et al.* (2000) reported that four of 52 adult males (from three litters) exposed perinatally to DINP exhibited malformations: one displayed a fluid-filled testis, a second displayed paired testicular and epididymal atrophy, the third displayed bilateral testicular atrophy and the fourth displayed unilateral epididymal agenesis with hypospermatogenesis and scrotal fluid-filled testis devoid of spermatids. The low incidence of reported effects was without any dose response, using a small number of rats, and effects are of unclear significance. The collective incidence of

effects in DINP treated animals was 7.7% (compared to 82% with DEHP treated animals). No endpoint on its own was significantly different from control values; rather, different effects were pooled to produce the 7.7% incidence. This type of data manipulation is not routinely performed in toxicological safety evaluations, nor is it considered good statistical practice. Based on the above points (historical control data and pooling of data to achieve significance), the significance of the reported findings is questionable.

Likewise, DINP does not induce general reproductive tract malformations manifested as decreased weights in androgen sensitive tissues: levator ani/bulbocavernosus muscles (LABC), seminal vesicles, ventral prostrate, glans penis, bulbouretral gland, and epididymis (Adamsson *et al.*, 2009; Boberg *et al.*, 2011; Gray *et al.*, 2000). These findings are not unexpected since, as discussed above, DINP only induces transient effects on fetal testosterone.

Some effects in androgen sensitive tissue weight were reported by Lee and Koo (2007) in a study similar in design to the Hershberger assay was utilized. However, DINP did not induce consistent changes in these androgen sensitive tissues. A significant decrease in seminal vesicle weight was observed in all DINP dose groups while a significant decrease in LABC weight was only observed in the high dose group. Regardless of control group, the weights of the sex accessory tissues from the administered groups showed no consistent or dose-related significant differences from the testosterone-only animals. In both of these cases, the data do not meet the Organisation for Economic Co-operation and Development (OECD) or Environmental Protection Agency (EPA) criteria for being classified as having a positive results since not all tissues were effected and no dose-response was observed.

The dietary study with DINP conducted at the Hamner Institutes (Clewell *et al.*, 2012b) included evaluation of phallus malformation, preputial separation, a full suite of reproductive organ weights at PND 49 and a comprehensive review of testes and epididymal histopathology at PND 2 and PND 49. Global endpoint analysis showed no evidence of a rat “phthalate syndrome” on PND 49 with DiNP administration (Clewell *et al.*, 2012b).

4.2.5 There is no Strong Evidence DINP Adversely Affects Sperm Production or Morphology

Two studies have examined sperm counts in male rats exposed to DINP (Boberg *et al.*, 2011; Kwack *et al.*, 2009). Boberg *et al.* (2011) reported that on post natal day 90, a small but significant ($p = 0.048$) *increase* in sperm count was observed in male offspring from dams that were exposed to 900 mg/kg/day DINP between gestation day 7 and post natal day 17; however, based on an increase in sperm counts measured as sperm per gram cauda epididymis and a slight decrease in epididymis weight, the authors concluded that “these data may indicate that DINP does not affect testicular sperm production”. Conversely, Kwack *et al.* (2009) reported a reduction in sperm count (~25%) in adult males exposed to 500 mg/kg/day DINP for 4-weeks

beginning at 28 days of age. However, Kwack *et al.* (2009) reported no effects on sperm quality or motility.

The reduction in sperm count observed in Kwack *et al.* (2009) is of questionable relevance since higher doses of DINP and DIDP were used in the definitive two-generation reproductive toxicity studies where no effects on fertility were reported in males that would have been exposed to each substance for a longer period of time, including both the P and F_1 generations. Fertility is dependent not only on having adequate sperm count, but also on having normal sperm quality. When sperm quality is good, (*i.e.* normal motility as demonstrated for DINP in Kwack *et al.* (2009)), then a significant reduction in sperm count is required to affect fertility (Parker, 2006). Furthermore, Kwack *et al.* (2009) did not assess reproductive performance in these animals, critical to the interpretation of their findings.

4.2.6 DINP Does Not Affect the Onset of Puberty or Male Mating Behavior

DINP exposure during gestation had no effect on the age of preputial separation in male rats (Gray *et al.*, 2000; Masutomi *et al.*, 2003). Furthermore, as reported by Lee *et al.* (2006a; 2006b), the frequency of copulatory behaviors in post natal week 20 animals was unaffected by DINP at doses of 400 or 4000 ppm (number of mountings, number of intromissions, number of ejaculations, and post ejaculation interval). These observations support the findings of the definitive two-generation reproductive and developmental toxicity study in which there are no adverse effects reported for male fertility parameters (Waterman *et al.*, 2000).

The Clewell et al (2012b) study included evaluation of preputial separation (PPS) at PND 49. The PPS score was not altered with any of the DINP treatments.

4.2.7 DINP Does Not Impair Fertility

DINP has not been shown to alter male fertility in laboratory animals in the definitive two-generation reproductive and developmental toxicity study (Hushka *et al.*, 2001; Waterman *et al.*, 2000).[25] Impaired fertility would be considered the decisive concern and ultimate result of the collective effects described for the male reproductive tract and termed “rat phthalate syndrome”. As previously described, there were no effects on male fertility parameters or reproductive performance in either the parental (P) or first filial (F_1) generation. These studies demonstrate that adult males (P) exposed to DINP prior to mating are successfully able to reproduce. More importantly, the reproductive capacity of the F_1 generation males that were exposed to both

[25] Conducted according to EPA Health Effects Test Guideline OPPTS 870.3800 and in accordance with the principles of Good Laboratory Practices.

chemicals throughout their lifetime was unaltered. Therefore, it is clear that DINP does not impair fertility[26].

4.3 Conclusion: DINP Does Not Induce "Rat Phthalate Syndrome"

There has been speculation or an assumption that the combination of phenomena associated with exposure to low molecular weight phthalates in laboratory rodents, "rat phthalate syndrome," can be extended to include high molecular weight phthalates and is relevant to humans. Proposed key events critical to the induction of the hypothesized "rat phthalate syndrome" include a decrease in fetal testosterone and insl3 (Gray and Foster, 2003; National Research Council, 2008). It is important to again emphasize that the mechanisms underlying these effects remain ill-defined.

A decrease in fetal testosterone levels has been observed in three studies with DINP (Boberg *et al.*, 2011; Borch *et al.*, 2004; Clewell *et al.* 2012a); however, it appears to be a transient effect (Boberg *et al.*, 2011; Clewell *et al.* 2012b). Furthermore, there is a strong disconnect between this observed hormone change and the absence of predicted adverse phenotypes. The most sensitive phenotypic endpoints for the identification of "rat phthalate syndrome" are decreased anogenital distance and nipple retention (Carruthers and Foster, 2005; Gray *et al.*, 2009; National Research Council, 2008; Wilson *et al.*, 2007). While Boberg *et al.* (2011) reported a significant decrease in anogenital distance in males gestationally exposed to DINP (900 mg/kg/day) on post natal day 13 (approximately 6%), there was no difference between treated animals and controls on post natal day 90; the effect was transitory. Additionally, there was no effect on nipple retention at either time point. The decrease observed in Boberg *et al.* also may not have been due to an anti-androgenic process. A similar decrease in AGD was seen in Clewell *et al.* (2012b) at PND 14, however no difference was found at any dose level, in the same animals, at either PND 2 or PND 49. Since effects on AGD occur *in utero*, the differences would be expected to be observed at PND2, not that they would appear later. In addition the measure is sensitive to size, and pups weights at the effect dose at that time point were statistically lower than controls. No effects on nipple retention were observed in the study at any time point or dose level. No effects on AGD or nipple retention were observed in the definitive two-generation reproductive toxicity test on DIDP (Hushka *et al.*, 2001).

[26] For its monographs of seven phthalates, the National Toxicology Program Center for Evaluation of Risks to Human Reproduction (NTP-CERHR) created a scale ranging from clear, some, or limited evidence of adverse effects to insufficient evidence for a conclusion to limited, some or clear evidence of no adverse effects. The conclusion with respect to reproductive toxicity was "limited evidence of no adverse effects" for DINP and "some evidence of no adverse effects" for DIDP. In contrast, the conclusion for BBP for male reproductive toxicity was "some evidence of adverse effects" and for DBP and DEHP the conclusion for reproductive toxicity was "clear evidence of adverse effects." The NTP-CERHR evaluations can be accessed at http://ntp.niehs.nih.gov/pubhealth/hat/noms/index.html.

Additionally, DINP has been shown not to induce hypospadias, cryptorchidism, or alter the androgen sensitive tissues. Furthermore, in the definitive two-generation reproductive toxicity tests, had no effect on fertility or developmental parameters.

Overwhelmingly, the data clearly indicate that DINP does not induce the adverse effects hypothesized to be part of "rat phthalate syndrome". Therefore, the applicability of the "syndrome" for hazard assessment is not supported. Limited research suggests that DINP induces a reduction in fetal testosterone synthesis. However, use of decreased testosterone as the sentinel event predictive of adverse effects is problematic as DINP does not induce the effects consistent with the hallmarks of the "rat phthalate syndrome." In addition, species specific differences in sensitivity to phthalate induced disruption in testosterone are clear. Recent and developing evidence indicates that humans are more similar to mice in that both seem to be refractory to phthalate induced testosterone reductions. Therefore, the relevance of this endpoint for human hazard or cumulative risk assessment is highly questionable.

Finally the conclusion that DINP does not induce "Rat Phthalate Syndrome" is supported by the summary of results tables provided in the CHAP report.

The report "defined male developmental and reproductive toxicity via an anti-androgenic mode of action as the critical effect," therefore "the CHAP deemed it as important to use such responses as the basis for cumulative risk assessments"(pg 61). It is important to note that when conducting a CRA all components of the CRA should share the same MOA. However, in the CHAP's own table, the key events in the mechanism of action proposed to lead to the adverse developmental effects observed in rodents are not consistent for DINP as shown in the figure below:

Summary of Mechanism of Action Studies									
Chemical	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
DBP	↓	↓		↓		↓	↓	↓	
BBP	↓	↓							
DEHP	↓	↓	↓	↓	↓	↓	↓	↓	↓
DEHP+DBP	↓	↓	↓	↓					
DNOP									
DINP	↓	↑	↓	↓	↑			↑	
DIDP									
DMP									
DEP									
DIBP	↓	↓		↓		↓		↓	↓
DPENP	↓	↓	↓	↓					
ATBC									
DEHA									
DINX									
DEHT									
TOTM									
TPIB									

Results from MoA studies for DINP are inconsistant with other phthalates used in the CRA

1 = Testosterone
2 = insl3 (Insulin-like factor 3)
3 = CYP11A (Rate-limiting enzyme responsible for the conversion of cholesterol to pregnenolone)
4 = StAR = Steroidogenic Acute Regulated Protein, involved in mitochondrial cholesterol uptake
5 = LH = Lutenizing Hormone
6 = SR-B1 = Scavenger Receptor B-1, responsible for cholesterol uptake by Leydig cells
7 = PBR = Peripheral Benzodiazepene Receptor, involved in mitochondrial cholesterol uptake
8 = CYP450scc = Cytochrome P450 side chain cleavage enzyme, steroid converting enzyme
9 = SF-1 = Nuclear Receptor Steroidogenic Factor-1, regulates expression of genes involved in steroidogenesis

One of the reviewers[27] (Reviewer 1) pointed out that there are at least three known MoA's operating to produce phthalate syndrome – "we know that there are at least three distinct MOAs operating to produce the syndrome of responses associated with *in utero* exposure to specific phthalates". However, this comment was not addressed by the CHAP. The lack of concordance in mechanism of action studies for DINP calls into question the inclusion of DINP in the cumulative risk assessment. It is of note that in the recent publication by EPA scientists on cumulative risk assessment of phthalates "the HI calculations do not include DINP since the critical effect is not in the reproductive/developmental domain." (Christensen et al 2014)

In a summary of effects considered to be the consequence of anti-androgenic processes (Table A-9 from CHAP report below), it is apparent that DINP causes few if any of the effects observed in studies of other phthalates. The lack of consistency between DINP and the other phthalates used in the cumulative risk assessment in both the mechanism of actions studies and the developmental toxicity demonstrates that inclusion of DINP in the cumulative risk assessment was not justified.

[27] Toxicology Excellence for Risk Assessment (TERA) arranged for written peer review of the draft CHAP report. As described in the peer review report: "The goal of the expert review was to provide CPSC and the Chronic Hazard Advisory Panel (CHAP) with independent scientific and technical expert opinion and comment on the draft text. The objective of the peer review was to obtain a broad, high-level peer review of the report, focusing on the overall risk assessment process that the CHAP applied to phthalates, and in particular on the novel methods the CHAP used (e.g., development of distributions of hazard indices for cumulative risk)."

Table A-9 Summary of animal male developmental toxicology.

PE	Testis malform. /histopathology	Testis wt.	Seminal vesicle	Epididymal wt.	Cryptorchidism	Hypospadias	Gubernacu-lar malformations
DBP	↑	↓	↓	↓	↑	↑	↑
BBP	↑	↓	↓	↓	↑	↑	↑
DEHP	↑	↓	↓	↓	↑	-	
DNOP							
DINP	-	↓	-	-			
DIDP							
DMP	-	-	-	-			
DEP	-	-	-	-	-	-	-
DIBP	↑	↓	↓?	↓	↑	↑	↑?
DPP	↑	↓		↓	↑?	↑?	↑?
DHEXP					↑		
DCHP					↑	↑	
DIOP							
DPHP							
ATBC							
DEHA		-	-	-			
DINCX					-?	-?	-?
DEHT							
TOTM							
TPIB							

↑= increase; ↓= decrease; - = not affected; PE = phthalate esters

5. Use of Animal Data to Assess Hazard and Risk

5.1 The Toxicological Data were used to Introduce Multiple Layers of Conservatism in the CRA

Although it was intended that the CRA be conservative, it was also envisioned that it be evidence-based, sound, and reasonable. However, the way in which the toxicological data were used in the analyses introduced additional elements of conservatism into the risk assessment process.

As one example, Reviewer 2 commented that "L677-679 is of *substantial* concern. This sentence essentially says that more reliance was put on positive studies of quality and implies that negative studies of quality were of lesser value."

The lines indicated by the reviewer remain in the final text (emphasis added): ""**What should be done when confronted with conflicting results of animal studies?** Consider the quality and relevance of the studies, experimental design in the context of standard protocols, route of exposure, power, and confounders. The conservative approach is to **rely on the study reporting adverse effects** unless there are compelling reasons to exclude the study, *i.e.*, considerations such as quality, design, execution or interpretation."

This statement indicates the general bias of the report and highlights the extremely conservative nature of all approaches taken by the CHAP. Other examples of conservatism in the report include, but are not limited to:

- "The CHAP decided to take the conservative approach and to recommend a NOAEL of 50 mg/kg-d for BBP." (pg 87)
- "Using a weight-of-evidence approach, the CHAP has conservatively set the NOAEL for DEHP at 5 mg/kg-d." (pg 90)
- "Taking a conservative approach, the CHAP assigns the NOAEL for DINP at 50 mg/kg-day." (pg 98)
- "Using the more conservative of the two NOAELs from the 2008 Saillenfait study, the CHAP assigns a NOAEL of 125 mg/kg-day for DIBP." (pg. 111)
- "our approach is conservative in that it tends slightly to overestimate dose." (Appendix E1 – 42)

Of note is the conservative nature for the point of departures used for 4 out of the 5 phthalates included in the cumulative risk assessment.

Additionally the section "4.2 Species Differences in Metabolism, Sensitivity, and Mechanism" is decidedly biased toward identifying and highlighting studies suggesting potential anti-androgenic effects in non-rats, while dismissing or discounting results that do not support anti-androgenic effects in non-rats. As such, it does not provide an objective weight of evidence consideration. In effect, the CHAP has set forth the hypothesis that phthalates have anti-androgenic activity in non-rats and have not objectively considered evidence that fails to support their hypothesis.

A more balanced evaluation of the non-rat anti-androgenicity assessment data would identify that the weight of evidence indicates at most equivocal potential for activity. Given the nature of the available data, mostly indicating a lack of anti-androgenic activity outside rats, the null

hypothesis ("Phthalates are not anti-androgenic in non-rats") appears more scientifically supportable.

One final statement in the section may mislead some readers: "The experimental findings in the rat and the marmoset show that neonatal exposure to certain phthalates suppresses testosterone synthesis in the testes. These observations are highly relevant considering the high phthalate exposures that may occur in some neonates." The authors have not clarified what is meant by high exposures for some neonates, but may be referring to situations were neonates are subject to medical interventions that may result in relatively higher exposures to DEHP. It would be inappropriate generalize that specific scenario because it occurs infrequently and in the context of medical treatment, not use of toys.

5.2 Errors in Reporting DINP Data are Carried Throughout the Report

Table 2.1 records a NOAEL of 50 mg/kg/d for increased nipple retention from the 2011 Boberg paper. Doses examined in Boberg et al were 300, 600, 750, and 900 mg/kg/d. Nipple retention was noted as statistically significant at 750 and 900 mg/kg/d giving a NOAEL of 600 mg/kg/d for this endpoint. The authors indicate a NOAEL of 300 mg/kg/d. Statistically significant changes in sperm motility were observed at 600 mg/kg/d.

Table A-8 (Appendix A – 36) lists the consensus NOAEL for DINP as 300 mg/kg-d. Table 2.1 (pg. 24) lists a NOAEL of 50 mg/kg-d for DINP and references Boberg et al (2011) with the endpoint of increased nipple retention. Boberg et al 2011 did not use a dose of 50 mg/kg-d. The lowest dose tested was 300 mg/kg-d. Though the overall NOAEL from Boberg was determined as 300 mg/kg-d the NOAEL for nipple retention was set at 600 mg/kg-d. The value of 50 mg/kg-d appears to be carried forward for use in Case 3 (Table 2.15). Table D-8 in Appendix D cites back to Table 2.1 as basis for PoD for DINP.

This inconsistency was pointed out by multiple reviewers:

Reviewer 2 – "A comparison of Table 2.15 POD to the consensus NOAELs in Appendix A reveals a discrepancy for DINP (table 2.15 says 50, whereas Appendix A, L723 says 300)."

Reviewer 4 – "NOAELs reported in Table 2.1 cannot be located in the Appendix (Table A-8 and A-10)"

Table D-8 in Appendix D cites back to Table 2.1 as basis for PoD for DINP. The information in Table 2.1 is not correct. Table 2.1 (pg. 24) lists a NOAEL of 50 mg/kg-d for DINP and references Boberg et al (2011) with the endpoint of increased nipple retention. Boberg et al 2011

did not use a dose of 50 mg/kg-d. The lowest dose tested was 300 mg/kg-d. The NOAEL for nipple retention was set at 600 mg/kg-d.

6. Evaluation of Human Evidence

1. In the CHAP report Executive Summary, CHAP states "it is important to note that the phthalates for which associations were reported were not always consistent and differed across publications." (Page 2). It was also stated that "In some cases, adverse effects in humans were associated with diethyl phthalate exposure, although diethyl phthalate does not cause the phthalate syndrome in rats." Despite these cautions and cited inconsistencies the report makes overly strong statements, e.g., "In conclusion, these studies provide the first human data linking prenatal phthalate exposure (specifically DEP, DBP and DEHP) with antiandrogenic effects in male offspring." (pg. 28)

    - Reviewer 2 also noted the overstated conclusions based on weak epidemiological evidence "L1024 is a recommendation based on the epi neurodevelopmental studies. First, it doesn't belong here. Secondly, it is far too strong based to the information discussed immediately above." The text referenced by the reviewer remains in the final report.

2. Study selection and endpoint identification are not transparent

    - Methods for the selection and review of literature are not provided. This is a major omission in the report. Without a systematic approach, studies are chosen arbitrarily at best and intentionally at worst (*i.e.*, only those that best support a particular position), resulting in a biased picture of the issue and leading reviewers to draw incorrect and/or incomplete conclusions.

3. Key epidemiologic studies presented in the report were not evaluated for their relative contributions to our understanding of potential health effects of phthalate exposure in light of their specific methodological limitations. To the contrary, the CHAP presented results of each study at face value and not in light of important methodological considerations necessary to accurately place findings into context or weigh their relative import.

4. Additional studies provided in Appendix C as supporting evidence of the developmental reproductive and neurobehavioral health effects do not support the key studies either due to their own methodological limitations, which are noted by CHAP (see below), or, in at least one study (Rais-Baharami et al, 2004) there were no effects.

    - Of the four studies selected as supporting pubertal development and gynecomastia, Lomenick *et al.* 2007 was a study of no finding (*i.e.*, no differences in phthalate concentrations between girls with precocious puberty and controls without), Colon *et al.* 2000[28] and Durmaz *et al.* 2010 "had important limitations in methods…" and

[28] Comments on the limitations of this study were outlined in a letter to the editor [McKee, Richard H. "Phthalate exposure and early thelarche." *Environmental health perspectives* 112.10 (2004): A541.]

therefore, per the CHAP "need to be interpreted very cautiously due to critical limitations". (p. C6-C7)

- In a follow up study of 21 14-16 year olds who had undergone extracorporeal membrane oxygenation (ECMO) as premature infants Rais-Baharami et al, 2004 reported "all had normal values" for all biometrics and laboratory panels evaluated, including genital measurements and markers of pubertal development.

5. CHAP disregarded their own statements made in the report, suggesting readers use caution when interpreting the results of studies and drawing conclusions (see above sub-bullet). In addition, CHAP also disregarded conflicting information, such as discrepant study results, presented in their report.

    - As one example, in the CHAP Report Executive Summary, CHAP states "it is important to note that the phthalates for which associations were reported were not always consistent and differed across publications." It is not clear how this "important" consideration was accounted for in their conclusions. (Page 2). They also stated "In some cases, adverse effects in humans were associated with diethyl phthalate exposure, although diethyl phthalate does not cause the phthalate syndrome in rats." This would seem to be a major conflict of information that would need to be considered, but the report is silent with regard to complex, often conflicting information.

6. "Most studies primarily focus on the association of maternal phthalate exposure with male reproductive tract developmental endpoints and neurodevelopmental outcomes." (pg. 2)

    - This is not an accurate statement, as relatively few epidemiologic studies examined developmental reproductive outcomes, constituting a shallow pool of evidence from which to draw conclusions with respect to reproductive developmental endpoints. The limited number of studies is further complicated by numerous limitations and variability of results. (See details below).

    - Although the neurodevelopmental literature is more extensive than the reproductive literature, conclusions are similarly limited by methodological weaknesses and widely inconsistent results (See details below).

7. In summary, given the paucity and weaknesses of the epidemiologic data as presented in the report, conclusions are premature and may be inaccurate.

6.1 Evidence for Phthalates and Male Reproductive Tract Developmental Outcomes

- The report states (pg. 28) that "although the results of these studies were not entirely consistent, they represent the first human data to assess potential risks of developmental

exposure to phthalates". The report goes on to say that "these studies provide the first human data linking prenatal phthalate exposure…with anti-androgenic effects in male offspring."

- These strong conclusions of male reproductive developmental disruption are based on (1) a small pool of data, with (2) inconsistent findings (Huang *et al.* did not find shortened AGD in boys, as well as differences in metabolites), that (3) evaluated only one endpoint (AGD) and/or reported effects in only one endpoint and (4) have extensive methodological limitations.

- The above statement made by CHAP, suggesting that the issue of inconsistency across study results is ameliorated by virtue of the selected studies being "among the first" to examine AGD is fallacious. New, <u>un-replicated</u> study findings have limited value, particularly in the context of equivocal mechanistic/toxicologic data. The contribution and significance of a novel finding(s) can only be known after being established by scientific replication, which for AGD is as yet insufficient.

- The report goes on to say that these results have "important relevance to the hypothesized testicular dysgenesis syndrome" (TDS) (pg. 28).

  - The validity is to be questioned regarding the CHAP linking un-replicated study results with a hypothetical syndrome in order to support the final conclusions made in their report

- The CHAP characterizes AGD as a "relevant marker for reproductive health outcomes" and cites three supporting studies, two of adult men and one of male babies/toddlers (pg. 28).

  - <u>AGD has yet to be fully established as a marker of reproductive health and function</u>, although these early data are suggestive. In light of the extreme paucity of research on this subject, the CHAP's unqualified, definitive statement is premature.

  - Before definitive statements can be made regarding the clinical significance of AGD the following limitations found in the current literature must be addressed: (1) lack of standardized measurement of AGD. Standard measurements and optimal measurement methods are essential; (2) Limited normative data. Normalized age-specific population data in different ethnic groups is necessary to understanding clinical significance; and (3) establishing gold standard by which to adjust for body size Liu *et al.* 2014).

  - Although the two studies among adult men reported correlations between AGD and other structural deficits (e.g. testicular volume, penile length), as well as measures of testicular functioning (e.g. sperm count, motility) and fertility, AGD was not correlated with serum hormone levels (T, LH, FSH). In addition, there was some inconsistency between the two studies in that Eisenberg *et al.* (2011) reported AGD being correlated with all genital measurements, including testicular volume, whereas Mediola *et al.* (2011) did not.

- CHAP makes statements beyond the data of the three above cited studies attempting to connect the findings of these studies to phthalate exposure. The CHAP states that the three studies of AGD and reproductive health "demonstrated that shortened AGD is associated with reproductive conditions that are similar to those observed in rats with the phthalate syndrome. This observation supports the use of human AGD as a relevant measure to assess the anti-androgenic mode of action of phthalates during fetal development." However, these statements are contrary to the conclusions of the cited papers.

    - The studies by Liu *et al.* (2014), Eisenberg *et al.* (2011) and Mediola *et al.* (2011) did not examine phthalates specifically or exposure to any other substances.

    - These studies did not address etiology – *i.e.* what factors/set of conditions, etc. led to reduced AGD. A different study design (longitudinal cohort) with accurate and appropriate measurement of prenatal phthalate exposure is required to make the statements found in the CHAP report. As the literature currently stands, etiology is unknown.

    - As stated in Mediola *et al.* (2011) " Whether shorter AGD in men reflects such dysgenesis and whether this is a consequence of fetal antiandrogen exposure are speculative".

    - In a recent paper, Liu *et al.* (2014) stated "animal findings (from experiments of EDCs and AGD) cannot be extrapolated smoothly to human studies. The following reasons were cited: (1) high to low dose extrapolation and (2) lack of ability to distinguish between in utero androgen action from postnatal growth and environmental impact after birth. Moreover, the authors notably state "the complex nature of endocrine disruption requires more caution in order to interpret EDC-related health effects".

    - Recent studies have implicated prenatal maternal smoking, maternal age and parity in AGD (Fowler *et al.* 2011; Barrett *et al.* 2014). These factors were not considered by the CHAP as other potential explanation for reduced AGD, but these studies suggest these other potential explanations may have etiologic relevance.

- Swan *et al.* 2005 and 2008 have been widely criticized for methodological and statistical limitations[29]

    - Regarding outcome: third trimester sampling (not relevant time period of exposure), spot v. first morning void (not an accurate reflection of exposure) and lack of adjustment for urine concentration significantly compromise exposure assessment

    - Regarding the primary endpoint of interest – AGD – is also significantly compromised because there is no standard protocol for this measurement, wide range of age at measurement (2 months to 3 years), lack of normative data to serve as

[29] One example of commentary to Swan *et al.* 2005: McEwen GN, Jr, Renner G. Validity of anogenital distance as a marker of in utero phthalate exposure. Environ Health Perspect. 2006;114:A19–20.

referent, inter-rater variability and technical difficulty associated with measuring AGD in a pediatric population.

In summary, the current literature on phthalate exposure and reproductive outcomes, including AGD, is mixed and inconclusive, despite the final conclusions made by the CHAP.

6.2 Evidence for Phthalates and Neurodevelopmental Outcomes

- The CHAP Report cites a total of 9 studies on neurodevelopmental outcomes, which they describe as heterogeneous. Studies differ by study population (age in particular), type of psychometric test used, set of phthalates studied, timing of urine sample, statistics.
- Of all the tests conducted across these studies, few associations were observed.
- Results were also inconsistent – even in studies that followed the same cohort.
    - Engel et al (2009) reported better motor performance among males with higher exposure to low molecular weight phthalates, while females were reported to have poorer orientation and alertness with increasing concentration of high molecular weight phthalates
    - Engel *et al.* (2010) noted that adverse outcomes were observed only for boys with higher prenatal exposure to low molecular weight phthalates – this gender difference is not stated in the report.
- Given the range of neurodevelopmental/behavioral domains examined across these studies, the range of findings is narrow and the clinical relevance is not clear.
- Study results are taken at face value – there is no consideration of study limitations.

7. Use of Urinary Metabolite Measurements to Estimate Human Exposure

7.1 Selection of Exposure Data

The biomonitoring data used (*i.e.*, from NHANES 2005–2006) are not the most recently available. There is no clear rationale why the older data was used. But the use of this dataset results in an over-estimation of exposure as the urinary metabolites of some phthalates have been decreasing since 2005-06.

Neither DINP, DIDP nor their metabolites were reported as having been investigated in either Sathyanarayana *et al.* 2008a or 2008b. In fact, Sathyanarayana *et al.* (2008b) specifically stated that DINP metabolites were not measured. The exposure information from Sathyanarayana *et al.* is reported to cover the range of 1999–2005 and urine samples were analyzed by the CDC for

metabolite levels. The CDC did not include DINP metabolites in its standard phthalate panel during that time. It is of interest to note that a limited subset of infants reported on in the CHAP from the SFF study include DINP metabolite data. This difference in sample number between DINP and the other phthalates used in the CRA is inconsistently noted. For example Table 2.5 notes concentrations of DEHP and DINP metabolites indicating a sample size of 291 infants for the CHAP/SFF study, however, in Table 2.7 when reported daily intake levels the report makes a subnote for DINP of 67 infants. More importantly when reporting on the HI calculations for the SFF infants the report fails to note how the DINP hazard quotient (HQ) was incorporated. Appendix D indicates a sample size of 251 for HI calculations for the infant SFF population. Given that it is not possible for all those values to contain a HQ for DINP it is unclear how these analyses can be used for a recommendation for DINP.

7.2 Reporting of Results

The CHAP report authors stated results that are speculative; caveating their speculations with "hedge" phrases such as "This suggests common uses and/or common sources of exposure..."; "... an individual exposed to elevated amounts of one of the high molecular weight phthalates is likely exposed to elevated amounts of the other high molecular weight phthalates, too"; and "... the correlations are low to moderate, which indicates that the variability of each phthalate (metabolite) in urine is influenced by more than just one exposure source and that exposures are similar."

The CHAP report authors concluded that exposures of pregnant women compared to non-pregnant women in the 15 to 45 age range were not statistically different. Yet, given the paucity of data on pregnant women in the NHANES data set, it is unclear why the CHAP used the data on pregnant women in preference to the much larger data set for women of child-bearing age.

The CHAP report authors speculate, but provide no evidence, that an individual exposed to elevated amounts of one of the high molecular weight phthalates will likely be exposed to elevated amounts of other high molecular weight phthalates as well.

7.3 Statement of Conclusions

Related to the first bullet in Section 2.5.4, many of the Conclusions seem inconclusive based on the use of caveats, e.g., "... indicating high exposures in some women"; "... there are indications of similar correlations"; "This suggests that sources and routes of exposure are similar..."; and "... we assume it is highly likely that the substitution of one phthalate will lead to increased exposure to another (similar) phthalate."

8. Use of Scenario-Based Exposure Assessment

8.1 Comments on Modeling of Phthalate Exposure from Food

8.1.1 Errors in use of food concentration data

The phthalate food commodities concentration data (CHAP report, Table E1-16) are high in comparison to other published assessments. As one example, the DINP and DIDP concentrations are several orders of magnitude greater than the estimates by Wormuth *et al.* (2006) and Clark *et al.* (2011).

Since the data were based on, "Bradley (2011) as described in Carlson and Patton (2012)," the Bradley report was obtained. It seems evident that the CHAP has misinterpreted the data in Bradley when calculating the phthalate concentrations shown in Table E1-16. The concentration data in Table E1-16 are in units of **µg/g**, however, the concentration data in Bradley (2011) are in units of **µg/kg**. Therefore, if the data from Bradley (2011) were not converted correctly, the concentration values in Table E1-16 (reproduced below) would be too large by a factor of 1000. Despite the error in the food concentration values in Table E1-16 it appears that the CHAP may have properly converted the concentration values prior to calculating dietary exposure. However, the lack of transparency in the use of the data inputs into the equation for estimating exposure from direct ingestion makes it difficult to know with certainty whether dietary exposure was calculated correctly or not. The uncertainty is compounded since only aggregate dietary exposure results are reported for each age group. The lack of transparency associated with the use of data in the CHAP exposure assessment makes for uncertain results that degrade any confidence in the accuracy of the exposure assessment. Given the uncertainty and lack of transparency in the dietary exposure assessment, the following concerns are noted:

- Table E1-16 should be updated to reflect the correct units for the food commodities concentrations, and that the values, with correct units, used in Table E1-16 are rechecked for accuracy

- A more transparent accounting of the calculation of dietary exposure should be provided. This would entail explicitly showing the calculations with data inputs used to calculate dietary exposure for at least one age group. The calculations should be shown separately for each food type and phthalate ester with the sum of these reflected in the aggregate dietary exposure value provided in the table in the Supplemental Data section of the report associated with the appropriate age group.

Table E1-16 Mean and 95th percentile concentrations of selected phthalate esters in food commodities (µg/g).[a]

Food Commodity		DEP	DBP	DIBP	BBP	DNOP	DEHP	DINP	DIDP
Grain	*Mean*	5.1	12.3	25.2	9.0	12	78	639	393
	0.95	11.4	35.4	91.6	25.7	35	234	2984	1198
Dairy	*Mean*	21.1	6.8	18.2	7.1	12	173	508	326
	0.95	89.2	17.2	69.9	16.4	26	554	1394	943
Fish	*Mean*	13.6	12.8	10.0	14.7	17	98	819	377
	0.95	40.2	51.5	40.7	46.6	45	286	2174	1281
Meat	*Mean*	5.1	6.8	5.5	12.2	11	54	298	236
	0.95	16.1	28.3	14.2	35.0	38	191	927	986
Fat	*Mean*	7.2	20.8	17.3	108.8	47	689	1481	1055
	0.95	29.2	54.2	46.5	93.2	133	2784	2851	2397
Eggs	*Mean*	4.7	5.2	5.7	9.4	20	24	385	259
	0.95	8.2	8.8	10.9	19.8	71	39	742	407

[a] Mean and 95th percentile concentrations were estimated from data in Bradley (2011) as described in Carlson and Patton (2012). Nondetects were treated as one-half the detection limit.

- P. E1-5, Eq. (5) – The dermal absorption rate of the PE of interest is based on the rate of dermal transfer and absorption of DEHP (given as 0.24 µg/cm^2-h (from Deisinger *et al.* 1998) multiplied by the ratio of the phthalate ester's (PE) percutaneous absorption rate to the percutaneous absorption rate for DEHP (these are given in Table E1-14). For DINP the resulting dermal absorption rate is given as **0.2 µg/cm^2-h**.

  However, Deisinger and co-workers calculated the dermal absorption rate for DEHP through rat skin. Elsisi *et al.* (1989) examined the extent of dermal absorption of a series of phthalate diesters in the rat. This study showed that DIDP was substantially less dermally absorbed (by about a factor of 10) than DEHP. Further, Scott *et al.* (1987) measured the absorption of undiluted phthalate diesters, including DEHP, *in vitro* through human and rat epidermal membranes. Scott and co-workers reported that the absolute rates of absorption measured indicated that the phthalate esters were slowly absorbed through both human and rat skin with human skin being approximately 4 times less permeable than rat skin.

  Based on the dermal absorption rate for DEHP of 0.24 µg/cm^2-h (Deisinger *et al.* 1998), an adjustment for decreased absorption of DINP compared to DEHP (Elsisi *et al.* 1989), and an adjustment factor for decreased absorption of human skin compared to rat skin (Scott *et al.* 1987), the resulting dermal absorption rate for DINP should be closer to **0.006 µg/cm^2-h**.

  The CHAP report dismisses this by stating that such adjustments were unnecessary since the permeability of human skin varies by anatomic site and that rodent skin *may be* an adequate model for neonatal skin because neonatal skin is more permeable than adult human skin.

However, it is highly unlikely to be 33 fold more permeable, as would be implied by the CHAP's value (0.2 µg/cm2-h) versus the value calculated from study data (0.006 µg/cm2-h.

- The exposure duration, *i.e.*, 24 hours, for dermal contact with Personal Care Products for women (Table E1-8), infants (Table E1-9), toddlers (Table E1-10), children (Table E1-11), and Household Products for women (Table E1-8) seems unrealistically high. This is a conservative assumption and likely results in overestimation of the true dermal exposure to phthalates.

- The dietary exposures were calculated by summing the contribution from each food category using Equation (1) (as stated on page E1-25). However, this will likely result in overestimation of dietary consumption of phthalates since it implies that all food categories are eaten every day. This is probably a conservative assumption

.

8.1.2 Inappropriate replacement of "nondetect" values with values equal to half the level of detection

- Section 2.4 – Food items with "nondetects" in the Food Item Phthalate Residue databases were assigned a value of one-half the Level of Detection (LOD) or one-half the Level of Quantification (LOQ). This was referred to as "replacement." The authors justified replacement on the assumption that residues were present in the food since, phthalates were ubiquitous in the environment, and therefore by extension, ubiquitous in food stuffs. Further, residues that were "not confirmed" in the dataset were treated as equivalent to the measured values and the replacement values.

  The result was that the majority of DINP dietary residues were replacement values and all of the DIDP residues were replacement values (see Section 3.2.1). In conclusion, the estimates of exposure for DINP and DIDP were not based on measured data but rather on the limits of the analytical methodology. Further, it is not clear what input values were used for the Consumption Factors (Appendix E3, Section 2.5.2) to account for the fraction of the population eating a specific food type, or the fraction of phthalate absorbed by the gastrointestinal tract (Appendix E3, Section 2.5.4). The resulting lack of transparency in the data used to calculate dietary exposure makes it difficult to replicate the dietary exposure estimates.

8.2 Methodology

8.2.1 Comment on Results

The CHAP report states that pregnant women were also exposed to DINP from the indoor environment. While this may be true, it should be noted that only about 5% of women's exposure to DINP was due to indoor environment. The overwhelming contribution was from diet (about 95%) (CHAP report, Fig. 2.1).

8.2.2 Comment on Design

The CHAP mentions that in order to complete the scenario-based exposure estimates, data prior to 2000, as well as professional judgment was used to estimate some of the inputs. The lack of reliable input data on people's habits, lifestyle, and product use made using default values in the analysis necessary for at least some of the exposure factors. The use of default values for exposure factors usually results in exposure estimates that are inflated compared to actual exposures. Despite this, the CHAP did not do a satisfactory job of identifying and quantifying the uncertainties in the scenario-based exposure analysis.

8.2.3 Comments on Conclusion

The CHAP states that, the findings of this study were more or less in agreement with other phthalate exposure assessments. They go on to state that, the CHAP's results are within an order of magnitude of other findings. Neither of these statements is correct – for adult females, the CHAP estimates were much greater than 1 order of magnitude than, for example, Wormuth *et al.* (2006). Further, the CHAP estimates for children were at least an order of magnitude greater than either Wormuth *et al.* or Clark *et al.* (2011).

9. The Recommendations

An inconsistency in recommendations was noted by Review 4 "For instance the permanent ban on DNOP is recommended to be lifted. There is little evidence of reproductive and development effects of DNOP, but evidence on systemic effects is strong. With respect to population exposure, the MOE is higher. In comparison, because of the evidence of reproductive and developmental effects, DIBP and DINP made to the permanent ban list despite a higher MOE (Table 2.17) and less than 1% median contribution to the overall HI of 5 phthalates."

9.1 Comments on the Criteria for Recommendations

The CPSIA directed the CHAP to undertake a comprehensive examination of phthalates and phthalate alternatives and to determine whether any (other than the already permanently banned DBP, BBP and DEHP) should be declared "banned hazardous substances." For its examination, the CHAP identified endocrine-related effects as the critical endpoints, and, based on these data conducted a cumulative risk assessment which assessed the overall risk of DBP, BBP, DEHP, DIBP and DINP. The CHAP found no evidence that DMP, DEP, DNOP and DIDP had

endocrine modulating properties, and these phthalates were not included in the cumulative risk assessment. The alternatives were considered separately but as the data were limited, phthalate alternatives were not assessed.

Taking the screening level cumulative risk assessment at face value, the risks presented by this group of phthalates to the population at large are relatively low; when the analysis was based on median exposure estimates, the hazard index values for the three cases (see more detail below) were approximately 0.2, well below the concern level of 1. To the extent that concerns were identified, these were confined to the most highly exposed segments of the population and associated almost exclusively with DEHP. More specifically, of the fractional hazard quotients comprising the hazard index, approximately 90% was associated with DEHP, 10% of the risk was associated with DBP, and the risks associated with the remaining phthalates (DIBP, BBP and DINP) were negligible. In the more extreme cases, the attributed risk is almost entirely associated with DEHP. These conclusions are in line with, but quantitatively different from, three previously published risk assessments (Benson, 2009; Kortenkamp and Faust, 2010; Christensen *et al*., 2014).

The CHAP recommended that no action be taken on the permanent bans of DBP, BBP and DEHP; that the interim ban on DINP be made permanent; and that the interim bans on DNOP and DIDP be lifted as there was no evidence that these substances affected endocrine-related properties. Some of these recommendations are consistent with the CHAP's scientific evaluation but others are not. As the assessment showed that there might be risks from exposure to DEHP and DBP, particularly to the most highly exposed individuals, the recommendation to take no action on (*i.e.*, to leave in place) the permanent bans on DEHP and DBP might be appropriate based on their scientific evaluation. However, the analysis provided no reason to continue the ban on BBP. With respect to DINP, the analysis showed that there were wide margins between exposures and potential effects. As indicated, the CHAP analysis demonstrates the contribution of DINP to the cumulative risk is negligible. Nevertheless, the CHAP recommended that the interim ban be made permanent because of that negligible potential contribution. Further, as discussed in more detail below, the cumulative risk methodology was screening level only, used outdated exposure information, and relied on an excessively conservative hazard assessment for DINP. When these factors are taken into account, there is no scientific basis for finding DINP in children's products would pose an unreasonable risk.

9.2 Comments on Recommendation for DINP

In this section the CHAP states: "Taking a conservative approach, the CHAP assigns the NOAEL for DINP at 50 mg/kg-day." (pg. 98 CHAP report) Despite the conservative nature of the determined NOAEL the CHAP goes on to conduct a "simple extrapolation" based on relative potencies from ED50 values reported by Hannas *et al (2011)*. This approach is problematic for several reasons: First, the adjustment uses a NOAEL of 5 mg/kg-day for DEHP as a starting point. A NOAEL is an artifact of dose spacing and is not reflective a true NOAEL. The use of a bench mark dose, which considers the entire dose response curve, is more appropriate for this type of adjustment. Second, this type of potency adjustment is appropriate when there is limited data for a substance. This situation is not the case for DINP which has a large set of substance

specific data to draw from. Available robust *in vivo* studies for DINP provide scientifically defensible points of departure, as provided in the conservative NOAEL assigned for DINP by the CHAP.

Of additional concern is the use of the "simple extrapolation" value (11.5 mg/kg/d) in generating the lower bound estimate for the margins of exposure (MOE) used for DINP. It is not appropriate to use this type of modeled data when sufficient *in vivo* experimental data are available. In fact the CHAP report identified a "conservative NOAEL" of 50 mg/kg/d. When the "conservative NOAEL" identified in the CHAP report is used instead of the modeled value to calculate the MOE for anti-androgenic effects, section 5.3.2.4.3.1, the MOE is larger:

> "In infants in the SFF study, the MOE for total exposure ranged from ~~640~~ [2800] to 42,000 using 95th percentile estimates of exposure. For pregnant women, the MOE for total DINP exposure ranged from ~~1000~~ [4500] to 68,000. Typically, MOEs exceeding 100–1000 are considered adequate for public health; however, the cumulative risk of DINP with other anti-androgens should also be considered." (pg. 99)

The appropriate and conservative lower bound estimate for the MOEs are well above the 100-1000 range considered adequate for public health and larger than the MOEs for system liver effects. "Using the NOAEL of 15 mg/kg-d for systemic toxicity, the MOE for infants ranged from 830 to 4,200. The MOE for women ranged from 1600 to 15,000" (pg. 99). Additionally the CHAP indicated liver toxicity was the most sensitive endpoint for both DINP and DIDP. "The NOAEL for liver toxicity for DINP (12 mg/kg-day), as for DIDP, is lower than the lowest NOAEL for anti-androgenic toxicity (50 mg/kg-day for MNGs)." (Pg. 98)

Given that "[t]ypically, MOEs exceeding 100–1000 are considered adequate for public health" and the most sensitive endpoint was identified as systemic liver effects, the same used in conclusion from the earlier CHAP report, it is unclear how the CHAP came to its recommendation for DINP.

9.3 Recommendation for DIDP

We agree with the CHAP'S recommendation that the current ban on DIDP be lifted. However, we question the recommendation that U.S. agencies responsible for dealing with DIDP exposures from food conduct "the necessary risk assessments". The modeled estimates in the CHAP report from food are high in comparison to other published assessments. Given the lack of transparency and uncertainty in the CHAP's assessment associated with these calculations, as discussed above, it is questionable what conclusions can be drawn from the modeled data.

10. Supporting Documentation

10.1 A Sathyanarayana, S., Calafat, A.M., Liu, F., Swan, S.H., 2008a. Maternal and infant urinary phthalate metabolite concentrations: Are they related? Environ Res 108, 413–418.

The papers by Sathyanarayana *et al.*, 2008a and 2008b (*i.e.*, the SFF study) did not report measuring DINP, DIDP, or their metabolites. In fact, in one of those papers it is specifically stated that the urinary metabolites of DINP were not measured. However, the CHAP report included urinary metabolite data for DINP and DIDP (see Tables 2.5, 2.6, and 2.7) in it analyses and cited the SFF studies. According to the description in the Methodology (Section 2.5.3), urinary metabolite data were provided to the CHAP by Shanna Swan. Therefore, nothing is known about the DINP or DIDP data used by the CHAP, or how the data was evaluated to arrive at the values given in the CHAP report. If data on DINP and DIDP was collected in the SFF study, it is not known why it was not cited in either of the Sathyanarayana *et al.* papers.

- The stated objective of the SFF study was to investigate the relationship between maternal and infant urinary phthalate metabolite concentrations. However, the study design was flawed in that it could not definitively satisfy that objective.
    - The study did not ascertain the sources of phthalate exposure. For example, the researchers did not measure home environmental phthalate or dietary phthalate levels. As a result the Discussion and Conclusions are peppered with phrases such as, “it is likely,” “may explain,” ”suggesting,” and, ”it may be.”
    - The sample size was relatively small, 116 male infants and 94 female infants. Further, the study population was not very representative of the general population; being comprised of 80% white mother/infant pairs, 90.5% of which had health insurance.

10.2 B Sathyanarayana, S., Karr, C.J., Lozano, P., Brown, E., Calafat, A.M., Liu, F., Swan, S.H., 2008b. Baby care products: Possible sources of infant phthalate exposure. Pediatrics 121, e260–268.

- This paper appears to use the same overall database (*i.e.*, SFFII) used in Sathyanarayana *et al.* (2008a), but with small but noticeable differences in sample sizes and participant demographics.
- This study seeks to investigate the relationship between phthalate metabolite concentrations in infant urine and maternal reported use of dermally applied infant care products. The most obvious limitation to the study objective based on the methodology

is that the researchers did not have data on phthalate concentration in the personal care products (PCPs).

- The urinary phthalate metabolite data were obtained from wet diapers brought in on the day of the infant study visit. Infant urine samples were obtained by squeezing the diaper and collecting urine in containers.

- There are a number of limitations in this study:

    - Because of the method for collecting urine, *i.e.*, by squeezing out a diaper that was brought from home, it was not possible for the researchers to know the time that the urine sample was actually collected.

    - The concentration of phthalates in the PCPs was unknown. Thus, the analyses and conclusions on PCP use and phthalate exposure from them were speculative. In addition, the researchers did not know how much or how often the PCPs were used on an infant.

    - The questions in the questionnaire were often not specific. For example, one question asked, "How many hours per day does your infant usually spend playing with or using the following?" Categories listed were soft plastic toys/teething rings and pacifiers. Based on this question it is not possible to determine whether the child actually used any of the objects listed the day of, or the day before, the urine sample was collected. Further, the question asked about toy use but did not ask how much of the time was spent mouthing/sucking on the toy versus just handling them.

References

Adamsson, A., Salonen, V., Paranko, J., and Toppari, J. (2009). Effects of maternal exposure to di-isononylphthalate (DINP) and 1,1-dichloro-2,2-bis(p-chlorophenyl)ethylene (p,p'-DDE) on steroidogenesis in the fetal rat testis and adrenal gland. *Reprod Toxicol* **28**, 66-74.

Adham, I. M., Emmen, J. M., and Engel, W. (2000). The role of the testicular factor INSL3 in establishing the gonadal position. *Mol Cell Endocrinol* **160**, 11-6.

Akahori, Y., Nakai, M., Yakabe, Y., Takatsuki, M., Mizutani, M., Matsuo, M., and Shimohigashi, Y. (2005). Two-step models to predict binding affinity of chemicals to the human estrogen receptor alpha by three-dimensional quantitative structure-activity relationships (3D-QSARs) using receptor-ligand docking simulation. *SAR QSAR Environ Res* **16**, 323-37.

Akahori, Y., Nakai, M., Yamasaki, K., Takatsuki, M., Shimohigashi, Y., and Ohtaki, M. (2008). Relationship between the results of in vitro receptor binding assay to human estrogen receptor alpha and in vivo uterotrophic assay: comparative study with 65 selected chemicals. *Toxicol In Vitro* **22**, 225-31.

Anderson, W., Castle, L., Hird, S., Jeffery, J., and Scotter, M. (2011). A 20 volunteer study using deuterium-labelled phthalates to determine the kinetics and fractional excretion of primary and secondary metabolites of di-2-ethylhexylphthalate (DEHP) and di-isononylphthalate (DINP) used as urinary biomarkers of exposure. *Food and Chemical Toxicology* 49:2022-2029.

Aubert, M. L., Begeot, M., Winiger, B. P., Morel, G., Sizonenko, P. C., and Dubois, P. M. (1985). Ontogeny of hypothalamic luteinizing hormone-releasing hormone (GnRH) and pituitary GnRH receptors in fetal and neonatal rats. *Endocrinology* **116**, 1565-76.

Aylward, L. L., Hays, S. M., and Kirman, C. R. (2011). Urinary DEHP metabolites and food fasting time in NHANES. A report prepared by Summit Toxicology.

Babich, M., Chen, SB., Greene, M., Kiss, C., Porter, W., Smith, T., Wind, M., and Zamula, W. (2004). Risk assessment of oral exposure to diisononyl phthalate from children's products. Regulatory Toxicology and Pharmacology 40:151-167.

Barrett ES *et al*. Evidence for sexually dimorphic associations between maternal characteristics and anogenital distance, a marker of reproductive development. Am J Epidemiol. 2014; 1:57-66

Becker, K., Goen, T., Seiwert, M., Conrad, A., Pick-Fuss, H., Muller, J., Wittassek, M., Schulz, C., and Kolossa-Gehring, M. (2009). GerES IV: phthalate metabolites and bisphenol A in urine of German children. *Int J Hyg Environ Health* **212**, 685-92.

Benson, R. (2009). Hazard to the developing male reproductive system from cumulative exposure to phthalate esters – dibutyl phthalate, diisobutyl phthalate, butylbenzyl phthalate, diethylhexyl phthalate, dipentyl phthalate, and diisononyl phthalate. Regulatory Toxicology and Pharmacology 53:90-101.

Benson, R. (2009). Hazard to the developing male reproductive system from cumulative exposure to phthalate esters--dibutyl phthalate, diisobutyl phthalate, butylbenzyl phthalate, diethylhexyl phthalate, dipentyl phthalate, and diisononyl phthalate. *Regul Toxicol Pharmacol* **53**, 90-101.

Berman, T., Hochner-Celnikier, D., Calafat, A. M., Needham, L. L., and Amitai, Y. (2009). Phthalate exposure among pregnant women in Jerusalem, Israel: Results of a pilot study. *Environment International* **35**, 353-357.

Boas, M., Frederiksen, H., Feldt-Rasmussen, U., Skakkebaek, N. E., Hegedus, L., Hilsted, L., Juul, A., and Main, K. M. (2010). Childhood exposure to phthalates: associations with thyroid function, insulin-like growth factor I, and growth. *Environ Health Perspect* **118**, 1458-64.

Boberg, J., Christiansen, S., Axelstad, M., Kledal, T., Vinggaard, A., Dalgaard, M., Nellemann, C., and Hass, U. (2011). Reproductive and behavioral effects of diisononyl phthalate (DINP) in perinatally exposed rats. Reproductive Toxicology 31:200-209.

Boeniger, M.F. *et al.*, 1993. Interpretation of urine results used to assess chemical exposure with emphasis on creatinine adjustments – a review. Am. Ind. Hyg. Assoc. J. 54, 615–627.

Borch, J., Ladefoged, O., Hass, U., and Vinggaard, A. M. (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.

Borgert, C. J., Quill, T. F., McCarty, L. S., and Mason, A. M. (2004). Can mode of action predict mixture toxicity for risk assessment? *Toxicol Appl Pharmacol* **201**, 85-96.

Breous, E., Wenzel, A., and Loos, U. (2005). The promoter of the human sodium/iodide symporter responds to certain phthalate plasticisers. *Mol Cell Endocrinol* **244**, 75-8.

Calafat, A. M., Wong, L. Y., Silva, M. J., Samandar, E., Preau, J. J., Jia, L. T., and Needham, L. L. (2011). Selecting Adequate Exposure Biomarkers of Diisononyl and Diisodecyl Phthalates: Data from the 2005-2006 National Health and Nutrition Examination Survey. *Environ Health Perspect* **119**, 50-55.

Carruthers, C. M., and Foster, P. M. (2005). Critical window of male reproductive tract development in rats following gestational exposure to di-n-butyl phthalate. *Birth Defects Res B Dev Reprod Toxicol* **74**, 277-85.

Centers for Disease Control and Prevention (2009). Fourth National Report on Human Exposure to Environmental Chemicals.

Centers for Disease Control and Prevention (2011). Update to the Fourth National Report on Human Exposure to Environmental Chemicals

CHAP (2014). Report to the U.S. Consumer Product Safety Commission by the Chronic Hazard Advisory Panel on Phthalates and Phthalate Alternatives. U.S. Consumer Product Safety Commission, Directorate for Health Sciences, Bethesda, MD. 20814.

Christensen, K., Makris, S., and Lorber, M. (2014). Generation of hazard indices for cumulative exposure to phthalates for use in cumulative risk assessments. Regulatory Toxicology and Pharmacology 69:380-389.

Clark, K., David, R. M., Guinn, R., Kramarz, K. W., Lampi, M., and Staples, C. A. (2011). Modelling human exposure to phthalate esters: A comparison of indirect and biomonitoring estimation methods. *Journal of Human and Ecological Risk Assessment* **In press**.

Clark, R. L., Antonello, J. M., Grossman, S. J., Wise, L. D., Anderson, C., Bagdon, W. J., Prahalada, S., MacDonald, J. S., and Robertson, R. T. (1990). External genitalia abnormalities in male rats exposed in utero to finasteride, a 5 alpha-reductase inhibitor. *Teratology* **42**, 91-100.

Clemens, L. G., Gladue, B. A., and Coniglio, L. P. (1978). Prenatal endogenous androgenic influences on masculine sexual behavior and genital morphology in male and female rats. *Horm Behav* **10**, 40-53.

Clewell, R. A., Sochaski, M., Edwards, K., Creasy, D. M., Willson, G., & Andersen, M. E. (2013a). Disposition of diisononyl phthalate and its effects on sexual development of the male fetus following repeated dosing in pregnant rats. Reproductive Toxicology, 35, 56-69. Available online 17 July 2012 (2012a).

Clewell, R. A., Thomas, A., Willson, G., Creasy, D. M., & Andersen, M. E. (2013b). A dose response study to assess effects after dietary administration of diisononyl phthalate (DINP) in gestation and lactation on male rat sexual development. Reproductive Toxicology, 35, 70-80. Available online 18 August 2012 (2012b).

CPSC (2001). Chronic hazard advisory panel on diisononyl phthalate (DINP). U.S. CPSC, Bethesda, MD. June 2001.

CPSC (2002). Updated risk assessment of oral exposure to diisononyl phthalate (DINP) in children's products. U.S. CPSC, Bethesda, MD. 20814, June, 2002.

David, R. M. (2000). Exposure to phthalate esters. Environ Health Perspect 108, A440.

Davison, J.M., Noble, M.C.B., 1981. Serial changes in 24-hour creatinine clearance during normal menstrual cycles and the first trimester of pregnancy. Br. J. Obstet. Gynaecol. 88, 10–17.

Dufau, M. L., Catt, K. J., and Tsuruhara, T. (1972). A sensitive gonadotropin responsive system: radioimmunoassay of testosterone production by the rat testis in vitro. Endocrinology 90, 1032-40.

Fowler PA. *et al.* Maternal cigarette smoking and effects on androgen action in male offspring: unexpected effects on second-trimester anogenital distance. J Clin Endocrinol Metab 2011;9:E1502-1506

Ghisari, M., and Bonefeld-Jorgensen, E. C. (2009). Effects of plasticizers and their mixtures on estrogen receptor and thyroid hormone functions. *Toxicology Letters* **189**, 67-77.

Gray, L. E., and Foster, P. (2003). Significance of experimental studies for assessing adverse effects of endocrine-disrupting chemicals. *Pure Applied Chemistry* **75**, 2125-2141.

Gray, L. E., Jr., Barlow, N. J., Howdeshell, K. L., Ostby, J. S., Furr, J. R., and Gray, C. L. (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*.

Gray, L., Ostby, J., Furr, J., Price, M., Veeramacheneni, D., and Parks, L. (2000). Perinatal exposure to phthalates DEHP, BBP and DINP but not DEP, DMP or DOTP alters sexual differentiation of the male rat. Toxicological Sciences 58:350-365.

Habert, R., and Picon, R. (1984). Testosterone, dihydrotestosterone and estradiol-17 beta levels in maternal and fetal plasma and in fetal testes in the rat. *J Steroid Biochem* **21**, 193-8.

Hallmark, N., Walker, M., McKinnell, C., Mahood, I. K., Scott, H., Bayne, R., Coutts, S., Anderson, R. A., Greig, I., Morris, K., and Sharpe, R. M. (2007). Effects of monobutyl and di(n-butyl) phthalate in vitro on steroidogenesis and Leydig cell aggregation in fetal testis explants from the rat: comparison with effects in vivo in the fetal rat and neonatal marmoset and in vitro in the human. *Environ Health Perspect* **115**, 390-6.

Hannas, B ., Lambright, C., Furr, J., Howdeshell, K., Wilson, V., and Gray, L. (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. Toxicological Sciences 123:206-216.

Harris, C. A., Henttu, P., Parker, M. G., and Sumpter, J. P. (1997). The estrogenic activity of phthalate esters in vitro. *Environ Health Perspect* **105**, 802-11.

Heger, N., Hall, S., Sandrof, M., Hensley, J., Johnson, K. J., Houseman, E., Gaido, K. W., and Boekelheide, K. (2010). Interspecies approach to the assessment of human susceptibility to phthalate-induced endocrine disruption. *The Toxicologist* **114**, 1973 (Abstract).

Heger, N., Hall, S., Sandrof, M., Hensley, J., Johnson, K. J., Houseman, E., Gaido, K. W., and Boekelheide, K. (2011). Interspecies approach to the assessment of human susceptibility to phthalate-induced endocrine disruption. *The Toxicologist* **120**, 2191 (Abstract).

Hellwig, J., Freudenberger, H., and Jackh, R. (1997). Differential prenatal toxicity of branched phthalate esters in rats. *Food Chem Toxicol* **35**, 501-12.

Hildenbrand, S., Wodarz, R., Gabrio, T., and Volland, G. (2009). Biomonitoring of the di(2-ethylhexyl) phthalate metabolites mono(2-ethyl-5-hydroxyhexyl) phthalate and mono(2-ethyl-5-oxohexyl) phthalate in children and adults during the course of time and seasons. *Int J Hyg Environ Health* **212**, 679-84.

Hood, R. D., and Miller, D. B. (2006). Maternally mediated effects on development. In Developmental and Reproductive Toxicology (R. D. Hood, ed. CRC Press, Boca Raton, FL.

Hotchkiss, A. K., Lambright, C. S., Ostby, J. S., Parks-Saldutti, L., Vandenbergh, J. G., and Gray, L. E., Jr. (2007). Prenatal testosterone exposure permanently masculinizes anogenital distance, nipple development, and reproductive tract morphology in female Sprague-Dawley rats. *Toxicol Sci* **96**, 335-45.

Howdeshell, K. L., Furr, J., Lambright, C. R., Rider, C. V., Wilson, V. S., and Gray, L. E., 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.

Howdeshell, K. L., Rider, C. V., Wilson, V. S., and Gray, L. E. (2008a). Mechanisms of action of phthalate esters, individually and in combination, to induce abnormal reproductive development in male laboratory rats. *Environmental Research* **108**, 168-176.

Howdeshell, K. L., Wilson, V. S., Furr, J., Lambright, C. R., Rider, C. V., Blystone, C. R., Hotchkiss, A. K., and Gray, L. E., Jr. (2008b). 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**, 153-65.

Hushka, L. J., Waterman, S. J., Keller, L. H., Trimmer, G. W., Freeman, J. J., Ambroso, J. L., Nicolich, M., and McKee, R. H. (2001). Two-generation reproduction studies in rats fed di-isodecyl phthalate. *Reprod Toxicol* **15**, 153-69.

Imperato-McGinley, J., Binienda, Z., Arthur, A., Mininberg, D. T., Vaughan, E. D., Jr., and Quimby, F. W. (1985). The development of a male pseudohermaphroditic rat using an inhibitor of the enzyme 5 alpha-reductase. *Endocrinology* **116**, 807-12.

Imperato-McGinley, J., Binienda, Z., Gedney, J., and Vaughan, E. D., Jr. (1986). Nipple differentiation in fetal male rats treated with an inhibitor of the enzyme 5 alpha-reductase: definition of a selective role for dihydrotestosterone. *Endocrinology* **118**, 132-7.

Jarfelt, K., Dalgaard, M., Hass, U., Borch, J., Jacobsen, H., and Ladefoged, O. (2005). Antiandrogenic effects in male rats perinatally exposed to a mixture of di(2-ethylhexyl) phthalate and di(2-ethylhexyl) adipate. *Reprod Toxicol* **19**, 505-15.

Johnson, K. J., Heger, N. E., & Boekelheide, K. (2012). Of mice and men (and rats): phthalate-induced fetal testis endocrine disruption is species-dependent. Toxicological Sciences, 129(2), 235-248.

Karbe, E., and Kerlin, R. L. (2002). Cystic degeneration/Spongiosis hepatis in rats. *Toxicol Pathol* **30**, 216-27.

Kavlock, R., Boekelheide, K., Chapin, R., Cunningham, M., Faustman, E., Foster, P., Golub, M., Henderson, R., Hinberg, Il, Little, R., Seed, J., Shea, K., Tabacova, S., Tyl, R., Williams, P., and Zacharewski, T. (2002). NTP Center for the Evaluation of Risks to Human Reproduction: phthalates expert panel report on the reproductive and developmental toxicity of di-isononyl phthalate. Reproductive Toxicology 16:679-708.

Klimisch, H. J., Andreae, M., and Tillmann, U. (1997). A systematic approach for evaluating the quality of experimental toxicological and ecotoxicological data. *Regul Toxicol Pharmacol* **25**, 1-5.

Koch, H. M., and Angerer, J. (2007). Di-iso-nonylphthalate (DINP) metabolites in human urine after a single oral dose of deuterium-labelled DINP. *Int J Hyg Environ Health* **210**, 9-19.

Koch, H. M., Wittassek, M., Bruning, T., Angerer, J., and Heudorf, U. (2011). Exposure to phthalates in 5-6 years old primary school starters in Germany-A human biomonitoring study and a cumulative risk assessment. *Int J Hyg Environ Health* **In Press**.

Kohn, M. C., Parham, F., Masten, S. A., Portier, C. J., Shelby, M. D., Brock, J. W., and Needham, L. L. (2000). Human exposure estimates for phthalates. *Environ Health Perspect* **108**, A440-2.

Kortenkamp, A., and Faust, M. (2010). Combined exposures to anti-androgenic chemicals: steps towards cumulate risk assessment. International Journal of Andrology 33:463-474.

Kratochwil, K. (1971). In vitro analysis of the hormonal basis for the sexual dimorphism in the embryonic development of the mouse mammary gland. *J Embryol Exp Morphol* **25**, 141-53.

Kratochwil, K. (1977). Development and loss of androgen responsiveness in the embryonic rudiment of the mouse mammary gland. *Dev Biol* **61**, 358-65.

Kratochwil, K. (1986). Tissue combination and organ culture studies in the development of the embryonic mammary gland. *Dev Biol (N Y 1985)* **4**, 315-33.

Kratochwil, K., and Schwartz, P. (1976). Tissue interaction in androgen response of embryonic mammary rudiment of mouse: identification of target tissue for testosterone. *Proc Natl Acad Sci U S A* **73**, 4041-4.

Kruger, T., Long, M., and Bonefeld-Jorgensen, E. C. (2008). Plastic components affect the activation of the aryl hydrocarbon and the androgen receptor. *Toxicology* **246**, 112-23.

Kwack, S. J., Kim, K. B., Kim, H. S., and Lee, B. M. (2009). Comparative toxicological evaluation of phthalate diesters and metabolites in Sprague-Dawley male rats for risk assessment. *J Toxicol Environ Health A* **72**, 1446-54.

LaKind, J. S., Goodman, M., & Naiman, D. Q. (2012). Use of NHANES data to link chemical exposures to chronic diseases: a cautionary tale. PloS one, 7(12), e51086.

Lambright, C. S., Furr, J., Cardon, M., Hannas, B., Bermudez, D., Wrench, N., Wilson, V., Foster, P., and Gray, L. E. (2011). Fetal phthalate screen: Assessment of several phthalate esters (PE) on fetal rodent testosterone (T) production and gene expression following *in utero* exposure. *The Toxicologist* **120: Suppl 2 (Abstract)**.

Lee, B. M., and Koo, H. J. (2007). Hershberger assay for antiandrogenic effects of phthalates. *J Toxicol Environ Health A* **70**, 1365-70.

Lee, C. Y., and Ryan, R. J. (1973). Interaction of ovarian receptors with human luteinizing hormone and human chorionic gonadotropin. *Biochemistry* **12**, 4609-15.

Lee, H. C., Ko, Y. G., Im, G. S., Chung, H. J., Seong, H. H., Chang, W. K., Yamanouchi, K., and Nishihara, M. (2006a). Effects of phthalate/adipate ester exposure during perinatal period on reproductive function after maturation in rats. *Journal of Animal Science and Technology* **48**, 651-662.

Lee, H. C., Yamanouchi, K., and Nishihara, M. (2006b). Effects of perinatal exposure to phthalate/adipate esters on hypothalamic gene expression and sexual behavior in rats. *J Reprod Dev* **52**, 343-52.

Lin, S., Ku, H. Y., Su, P. H., Chen, J. W., Huang, P. C., Angerer, J., and Wang, S. L. (2011). Phthalate exposure in pregnant women and their children in central Taiwan. *Chemosphere* **82**, 947-955.

Liu C *et al.* Anogenital distance and its application in environmental health research. Environ Sci Pollut Res 2014; 21:5457-5464

Livera, G., Delbes, G., Pairault, C., Rouiller-Fabre, V., and Habert, R. (2006). Organotypic culture, a powerful model for studying rat and mouse fetal testis development. Cell and Tissue Research 324, 507-521.

Lohsiriwat, S., Imrittha, N., 2008. Effect of posture on creatinine clearance in late pregnancy and after pregnancy. J. Obstet. Gynaecol. Res. 34, 337–342.

MacSween, R. M. N., Burt, A. D., Portmann, B. C., Ishak, K. G., Scheurer, P. J., Anthony, P. P., and Weisenberg, E. (2003). *Pathology of the Liver, 4th edition*.

Martino-Andrade, A. J., Morais, R. N., Botelho, G. G., Muller, G., Grande, S. W., Carpentieri, G. B., Leao, G. M., and Dalsenter, P. R. (2008). Coadministration of active phthalates results in disruption of foetal testicular function in rats. *Int J Androl*. 32:704-712.

Masutomi, N., Shibutani, M., Takagi, H., Uneyama, C., Lee, K. Y., and Hirose, M. (2004). Alteration of pituitary hormone-immunoreactive cell populations in rat offspring after maternal dietary exposure to endocrine-active chemicals. *Arch Toxicol* **78**, 232-40.

Masutomi, N., Shibutani, M., Takagi, H., Uneyama, C., Takahashi, N., and Hirose, M. (2003). Impact of dietary exposure to methoxychlor, genistein, or diisononyl phthalate during the perinatal period on the development of the rat endocrine/reproductive systems in later life. *Toxicology* **192**, 149-70.

McKee, R. H., El-Hawari, M., Stoltz, M., Pallas, F., and Lington, A. W. (2002). Absorption, disposition and metabolism of di-isononyl phthalate (DINP) in F-344 rats. *J Appl Toxicol* **22**, 293-302.

Meek, M. E., Boobis, A. R., Crofton, K. M., Heinemeyer, G., Van Raaij, M., & Vickers, C. (2011). Risk assessment of combined exposure to multiple chemicals: A WHO/IPCS framework. Regul Toxicol Pharmacol, 60(2 suppl 1), S1-S14.

Mlynarcikova, A., Fickova, M., and Scsukova, S. (2007). The effects of selected phenol and phthalate derivatives on steroid hormone production by cultured porcine granulosa cells. *Altern Lab Anim* **35**, 71-7.

National Research Council (2008). *Phthalates and cumulative risk: The task ahead*. National Academies Press, Washington, DC.

Nef, S., and Parada, L. F. (1999). Cryptorchidism in mice mutant for Insl3. *Nat Genet* **22**, 295-9.

Ostby, J., Furr, J., and Gray, E. L. (2001a). The cumulative effects of the fungicide procymidone and the plasticizer di(n-butyl) phthalate on sexual differentiation in male Sprague-Dawley rats. *Biology of Reproduction* **64**, 347-348.

Ostby, J. S., Hotchkiss, A. K., Furr, R., and Gray, L. E. (2001b). Investigation of the ability of diisononyl phthalate (DINP) to alter androgen-dependent tissue development in Sprague-Dawley rats. *The Toxicologist* **60 (Abstract)**.

Parker, R. M. (2006). Testing for Reproductive Toxicity. In Developmental and Reproductive Toxicology (R. D. Hood, ed.), CRC Press, Boca Raton, FL.

Preau, J. L., Jr., Wong, L. Y., Silva, M. J., Needham, L. L., and Calafat, A. M. (2010). Variability over 1 week in the urinary concentrations of metabolites of diethyl phthalate and di(2-ethylhexyl) phthalate among eight adults: an observational study. *Environ Health Perspect* **118**, 1748-54.

Price, P. S., & Han, X. (2011). Maximum cumulative ratio (MCR) as a tool for assessing the value of performing a cumulative risk assessment. International journal of environmental research and public health, 8(6), 2212-2225.

Price, P., Han, X., Junghans, M., Kunz, P., Watts, C., & Leverett, D. (2012a). An application of a decision tree for assessing effects from exposures to multiple substances to the assessment of human and ecological effects from combined exposures to chemicals observed in surface waters and waste water effluents. Environmental Sciences Europe, 24(1), 1-13.

Price, P., Dhein, E., Hamer, M., Han, X., Heneweer, M., Junghan, M., & Kunz, P. (2012b). A decision tree for assessing effects from exposures to multiple substances. Environmental Sciences Europe, 24:26.

Remer, T., Neubert, A., and Maser-Gluth, C. (2002). Anthropometry-based reference values for 24-h urinary creatinine excretion during growth and their use in endocrine and nutritional research. *Am J Clin Nutr* **75**, 561-569.

Rider, C. V., Furr, J., Wilson, V. S., and Gray, L. E., Jr. (2008). A mixture of seven antiandrogens induces reproductive malformations in rats. *Int J Androl* **31**, 249-62.

Rider, C. V., Wilson, V. S., Howdeshell, K. L., Hotchkiss, A. K., Furr, J. R., Lambright, C. R., and Gray, L. E., Jr. (2009). Cumulative effects of in utero administration of mixtures of "antiandrogens" on male rat reproductive development. *Toxicol Pathol* **37**, 100-13.

Sathyanarayana, S., Calafat, A.M., Liu, F., Swan, S.H., (2008a). Maternal and infant urinary phthalate metabolite concentrations: Are they related? Environ Res 108, 413–418.

Sathyanarayana, S., Karr, C., Lozano, P., Brown, E., Calafat, A., Liu, F., and Swan, S. (2008b). Baby care products: Possible sources of infant phthalate exposure. Pediatrics 121:e260-e268.

Scott, H. M., Hutchison, G. R., Mahood, I. K., Hallmark, N., Welsh, M., De Gendt, K., ... & Sharpe, R. M. (2007). Role of androgens in fetal testis development and dysgenesis. Endocrinology, 148(5), 2027-2036.

Scott, H. M., Mason, J. I., and Sharpe, R. M. (2009). Steroidogenesis in the fetal testis and its susceptibility to disruption by exogenous compounds. *Endocr Rev* **30**, 883-925.

Su, Q., Benner, A., Hofmann, W. J., Otto, G., Pichlmayr, R., and Bannasch, P. (1997). Human hepatic preneoplasia: phenotypes and proliferation kinetics of foci and nodules of altered hepatocytes and their relationship to liver cell dysplasia. *Virchows Arch* **431**, 391-406.

Suzuki, Y., Niwa, M., Yoshinaga, J., Watanabe, C., Mizumoto, Y., Serizawa, S., and Shiraishi, H. (2009). Exposure assessment of phthalate esters in Japanese pregnant women by using urinary metabolite analysis. *Environ Health Prev Med* **14**, 180-7.

Takagi, H., Shibutani, M., Lee, K. Y., Masutomi, N., Fujita, H., Inoue, K., Mitsumori, K., and Hirose, M. (2005). Impact of maternal dietary exposure to endocrine-acting chemicals on progesterone receptor expression in microdissected hypothalamic medial preoptic areas of rat offspring. *Toxicol Appl Pharmacol* **208**, 127-36.

Takeuchi, S., Iida, M., Kobayashi, S., Jin, K., Matsuda, T., and Kojima, H. (2005). Differential effects of phthalate esters on transcriptional activities via human estrogen receptors alpha and beta, and androgen receptor. *Toxicology* **210**, 223-33.

Tietz, N. W. (2006). *Tietz Textbook of Clinical Chemistry and Molecular Diagnostics (4th. edition)*. Elsevier Saunders.

United States Consumer Products Safety Commission (2002). Response to Petition HP 99-1: Request to ban PVC in toys and other products intended for children five years of age and under. Submitted to the Commission with vote sheet on Sept. 18, 2002.

United States Environmental Protection Agency (1997). Exposure Factors Handbook.

United States Environmental Protection Agency (2002). Guidance on cumulative risk assessment of pesticide chemicals that have a common mechanism of toxicity.

Warren, D. W., Haltmeyer, G. C., and Eik-Nes, K. B. (1972). Synthesis and metabolism of testosterone in the fetal rat testis. *Biol Reprod* **7**, 94-9.

Waterman, S. J., Ambroso, J. L., Keller, L. H., Trimmer, G. W., Nikiforov, A. I., and Harris, S. B. (1999). Developmental toxicity of di-isodecyl and di-isononyl phthalates in rats. Reprod Toxicol 13, 131-6

Waterman, S., Keller, L., Trimmer, G., Freeman, J., Nikiforov, A., Harris, S., Nicolich, M., and McKee, R. (2000). Two-generation reproduction study in rats given di-isononyl phthalate in the diet. Reproductive Toxicology 14:21-36.

Wenzel, A., Franz, C., Breous, E., and Loos, U. (2005). Modulation of iodide uptake by dialkyl phthalate plasticisers in FRTL-5 rat thyroid follicular cells. *Mol Cell Endocrinol* **244**, 63-71.

Wilson, V. S., Howdeshell, K. L., Lambright, C. S., Furr, J., and Earl Gray, L., Jr. (2007). Differential expression of the phthalate syndrome in male Sprague-Dawley and Wistar rats after in utero DEHP exposure. *Toxicol Lett* **170**, 177-84.

Wittassek, M., and Angerer, J. (2008). Phthalates: metabolism and exposure. *Int J Androl* **31**, 131-8.

Wittassek, M., Koch, H. M., Angerer, J., and Bruning, T. (2010). Assessing exposure to phthalates - The human biomonitoring approach. *Mol Nutr Food Res* **54**, 1-25.

Wittassek, M., Wiesmuller, G. A., Koch, H. M., Eckard, R., Dobler, L., Muller, J., Angerer, J., and Schluter, C. (2007). Internal phthalate exposure over the last two decades--a retrospective human biomonitoring study. *Int J Hyg Environ Health* **210**, 319-33.

Ye, C. W., Gao, J., Yang, C., Liu, X. J., Li, X. J., and Pan, S. Y. (2009). Development and application of an SPME/GC method for the determination of trace phthalates in beer using a calix[6]arene fiber. *Anal Chim Acta* **641**, 64-74.

Ye, X., Pierik, F. H., Hauser, R., Duty, S., Angerer, J., Park, M. M., Burdorf, A., Hofman, A., Jaddoe, V. W., Mackenbach, J. P., Steegers, E. A., Tiemeier, H., and Longnecker, M. P. (2008). Urinary metabolite concentrations of organophosphorous pesticides, bisphenol A, and phthalates among pregnant women in Rotterdam, the Netherlands: the Generation R study. *Environ Res* **108**, 260-7.

Zacharewski, T. R., Meek, M. D., Clemons, J. H., Wu, Z. F., Fielden, M. R., and Matthews, J. B. (1998). Examination of the in vitro and in vivo estrogenic activities of eight commercial phthalate esters. *Toxicol Sci* **46**, 282-93.

Zimmermann, S., Steding, G., Emmen, J. M., Brinkmann, A. O., Nayernia, K., Holstein, A. F., Engel, W., and Adham, I. M. (1999). Targeted disruption of the Insl3 gene causes bilateral cryptorchidism. *Mol Endocrinol* **13**, 681-91.

Appendix A:
Percentiles for Weighted Daily Intake Values used in HI

September 2014

Appendix A: Percentiles for Weighted Daily Intake Values used in HI Calculations

Women of Reproductive Age (15-45)

NHANES 2005-2006

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: RIAGENDR = 2 AND RIDAGEYR > 14 AND RIDAGEYR < 46
by: Variable, SUDAAN Reserved Variable One, Percentiles.

for: Variable = DINP (μg/kg/day).

SUDAAN Reserved Variable One	Percentiles	Sample Size	Quantile	Std. Error
Total	50.00	618	0.95	0.07
	95.00	618	9.83	1.93
	99.00	618	25.85	9.98
1	50.00	618	0.95	0.07
	95.00	618	9.83	1.93
	99.00	618	25.85	9.98

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: RIAGENDR = 2 AND RIDAGEYR > 14 AND RIDAGEYR < 46
by: Variable, SUDAAN Reserved Variable One, Percentiles.

for: Variable = DEHP (μg/kg/day).

SUDAAN Reserved Variable One	Percentiles	Sample Size	Quantile	Std. Error
Total	50.00	618	3.50	0.27
	95.00	618	28.51	7.67
	99.00	618	205.17	.
1	50.00	618	3.50	0.27
	95.00	618	28.51	7.67
	99.00	618	205.17	.

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: RIAGENDR = 2 AND RIDAGEYR > 14 AND RIDAGEYR < 46
by: Variable, SUDAAN Reserved Variable One, Percentiles.

for: Variable = DBP (μg/kg/day).

SUDAAN Reserved Variable One	Percentiles	Sample Size	Quantile	Std. Error
Total	50.00	618	0.59	0.04
	95.00	618	2.79	0.40
	99.00	618	5.65	.
1	50.00	618	0.59	0.04
	95.00	618	2.79	0.40
	99.00	618	5.65	.

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: RIAGENDR = 2 AND RIDAGEYR > 14 AND RIDAGEYR < 46
by: Variable, SUDAAN Reserved Variable One, Percentiles.

for: Variable = DiBP (μg/kg/day).

SUDAAN Reserved Variable One	Percentiles	Sample Size	Quantile	Std. Error
Total	50.00	618	0.17	0.01
	95.00	618	0.89	0.20
	99.00	618	2.07	.
1	50.00	618	0.17	0.01
	95.00	618	0.89	0.20
	99.00	618	2.07	.

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: RIAGENDR = 2 AND RIDAGEYR > 14 AND RIDAGEYR < 46
by: Variable, SUDAAN Reserved Variable One, Percentiles.

for: Variable = BBP (μg/kg/day).

SUDAAN Reserved Variable One	Percentiles	Sample Size	Quantile	Std. Error
Total	50.00	618	0.22	0.01
	95.00	618	1.14	0.11
	99.00	618	2.43	0.66
1	50.00	618	0.22	0.01
	95.00	618	1.14	0.11
	99.00	618	2.43	0.66

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: RIAGENDR = 2 AND RIDAGEYR > 14 AND RIDAGEYR < 46
by: Variable, SUDAAN Reserved Variable One, Percentiles.

for: Variable = DINP (μg/kg/day).

SUDAAN Reserved Variable One	Percentiles	Sample Size	Quantile	Std. Error
Total	50.00	516	1.37	0.13
	95.00	516	11.79	1.68
	99.00	516	39.00	.
1	50.00	516	1.37	0.13
	95.00	516	11.79	1.68
	99.00	516	39.00	.

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: RIAGENDR = 2 AND RIDAGEYR > 14 AND RIDAGEYR < 46
by: Variable, SUDAAN Reserved Variable One, Percentiles.

for: Variable = DEHP (μg/kg/day).

SUDAAN Reserved Variable One	Percentiles	Sample Size	Quantile	Std. Error
Total	50.00	516	3.64	0.35
	95.00	516	32.54	6.81
	99.00	516	94.36	.
1	50.00	516	3.64	0.35
	95.00	516	32.54	6.81
	99.00	516	94.36	.

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: RIAGENDR = 2 AND RIDAGEYR > 14 AND RIDAGEYR < 46
by: Variable, SUDAAN Reserved Variable One, Percentiles.

for: Variable = DBP (μg/kg/day).

SUDAAN Reserved Variable One	Percentiles	Sample Size	Quantile	Std. Error
Total	50.00	516	0.67	0.07
	95.00	516	2.45	0.32
	99.00	516	5.14	.
1	50.00	516	0.67	0.07
	95.00	516	2.45	0.32
	99.00	516	5.14	.

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: RIAGENDR = 2 AND RIDAGEYR > 14 AND RIDAGEYR < 46
by: Variable, SUDAAN Reserved Variable One, Percentiles.

for: Variable = DiBP (μg/kg/day).

SUDAAN Reserved Variable One	Percentiles	Sample Size	Quantile	Std. Error
Total	50.00	516	0.27	0.01
	95.00	516	0.94	0.07
	99.00	516	2.26	0.48
1	50.00	516	0.27	0.01
	95.00	516	0.94	0.07
	99.00	516	2.26	0.48

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: RIAGENDR = 2 AND RIDAGEYR > 14 AND RIDAGEYR < 46
by: Variable, SUDAAN Reserved Variable One, Percentiles.

for: Variable = BBP (μg/kg/day).

SUDAAN Reserved Variable One	Percentiles	Sample Size	Quantile	Std. Error
Total	50.00	516	0.38	0.03
	95.00	516	1.67	0.27
	99.00	516	4.44	.
1	50.00	516	0.38	0.03
	95.00	516	1.67	0.27
	99.00	516	4.44	.

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: RIAGENDR = 2 AND RIDAGEYR > 14 AND RIDAGEYR < 46
by: Variable, SUDAAN Reserved Variable One, Percentiles.

for: Variable = DINP (μg/kg/day).

SUDAAN Reserved Variable One	Percentiles	Sample Size	Quantile	Std. Error
Total	50.00	568	2.73	0.30
	95.00	568	34.32	7.48
	99.00	568	112.70	.
1	50.00	568	2.73	0.30
	95.00	568	34.32	7.48
	99.00	568	112.70	.

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: RIAGENDR = 2 AND RIDAGEYR > 14 AND RIDAGEYR < 46
by: Variable, SUDAAN Reserved Variable One, Percentiles.

for: Variable = DEHP (μg/kg/day).

SUDAAN Reserved Variable One	Percentiles	Sample Size	Quantile	Std. Error
Total	50.00	568	1.97	0.15
	95.00	568	9.56	2.71
	99.00	568	54.87	.
1	50.00	568	1.97	0.15
	95.00	568	9.56	2.71
	99.00	568	54.87	.

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: RIAGENDR = 2 AND RIDAGEYR > 14 AND RIDAGEYR < 46
by: Variable, SUDAAN Reserved Variable One, Percentiles.

for: Variable = DBP (μg/kg/day).

SUDAAN Reserved Variable One	Percentiles	Sample Size	Quantile	Std. Error
Total	50.00	568	0.56	0.03
	95.00	568	2.10	0.57
	99.00	568	11.51	.
1	50.00	568	0.56	0.03
	95.00	568	2.10	0.57
	99.00	568	11.51	.

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: RIAGENDR = 2 AND RIDAGEYR > 14 AND RIDAGEYR < 46
by: Variable, SUDAAN Reserved Variable One, Percentiles.

for: Variable = DiBP (μg/kg/day).

SUDAAN Reserved Variable One	Percentiles	Sample Size	Quantile	Std. Error
Total	50.00	568	0.32	0.02
	95.00	568	0.94	0.09
	99.00	568	2.12	.
1	50.00	568	0.32	0.02
	95.00	568	0.94	0.09
	99.00	568	2.12	.

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: RIAGENDR = 2 AND RIDAGEYR > 14 AND RIDAGEYR < 46
by: Variable, SUDAAN Reserved Variable One, Percentiles.

for: Variable = BBP (μg/kg/day).

SUDAAN Reserved Variable One	Percentiles	Sample Size	Quantile	Std. Error
Total	50.00	568	0.21	0.02
	95.00	568	0.96	0.10
	99.00	568	2.01	0.26
1	50.00	568	0.21	0.02
	95.00	568	0.96	0.10
	99.00	568	2.01	0.26

NHANES 2011-2012

These data were retrieved before CDC removed the phthalate urinary metabolite data from the website in Aug 2014. The NHANES error in the data was with the weighting. These data did have weighting applied therefore the percentiles may be over or under estimated.

SUDAAN

Date: 08-11-2014 Page: 1
Time: 17:07:19 Table: 1

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: RIAGENDR = 2 AND RIDAGEYR > 14 AND RIDAGEYR < 46
by: Variable, SUDAAN Reserved Variable One, Percentiles.

for: Variable = DINP (μg/kg/day).

SUDAAN Reserved Variable One	Percentiles	Sample Size	Quantile	Std. Error
Total	50.00	461	4.40	0.75
	95.00	461	51.04	8.96
	99.00	461	140.97	.
1	50.00	461	4.40	0.75
	95.00	461	51.04	8.96
	99.00	461	140.97	.

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: RIAGENDR = 2 AND RIDAGEYR > 14 AND RIDAGEYR < 46
by: Variable, SUDAAN Reserved Variable One, Percentiles.

for: Variable = DEHP (μg/kg/day).

SUDAAN Reserved Variable One	Percentiles	Sample Size	Quantile	Std. Error
Total	50.00	461	1.53	0.10
	95.00	461	6.39	1.01
	99.00	461	15.81	.
1	50.00	461	1.53	0.10
	95.00	461	6.39	1.01
	99.00	461	15.81	.

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: RIAGENDR = 2 AND RIDAGEYR > 14 AND RIDAGEYR < 46
by: Variable, SUDAAN Reserved Variable One, Percentiles.

for: Variable = DBP (μg/kg/day).

SUDAAN Reserved Variable One	Percentiles	Sample Size	Quantile	Std. Error
Total	50.00	461	0.28	0.03
	95.00	461	1.36	0.10
	99.00	461	6.22	.
1	50.00	461	0.28	0.03
	95.00	461	1.36	0.10
	99.00	461	6.22	.

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: RIAGENDR = 2 AND RIDAGEYR > 14 AND RIDAGEYR < 46
by: Variable, SUDAAN Reserved Variable One, Percentiles.

for: Variable = DiBP (μg/kg/day).

SUDAAN Reserved Variable One	Percentiles	Sample Size	Quantile	Std. Error
Total	50.00	461	0.23	0.02
	95.00	461	0.88	0.09
	99.00	461	1.64	0.19
1	50.00	461	0.23	0.02
	95.00	461	0.88	0.09
	99.00	461	1.64	0.19

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: RIAGENDR = 2 AND RIDAGEYR > 14 AND RIDAGEYR < 46
by: Variable, SUDAAN Reserved Variable One, Percentiles.

for: Variable = BBP (μg/kg/day).

SUDAAN Reserved Variable One	Percentiles	Sample Size	Quantile	Std. Error
Total	50.00	461	0.17	0.01
	95.00	461	0.79	0.09
	99.00	461	1.71	.
1	50.00	461	0.17	0.01
	95.00	461	0.79	0.09
	99.00	461	1.71	.

Pregnant Women

NHANES 2005-2006

NHANES Survey Data 2005-2006 (drop zero daily intake data)
Females Ages 15 to 45

The SURVEYMEANS Procedure

Table 3

Data Summary	
Number of Strata	15
Number of Clusters	30
Number of Observations	2085
Number of Observations Used	628
Number of Obs with Nonpositive Weights	1457
Sum of Weights	63796631.8

Table 4

Statistics		
Variable	**Label**	**N Miss**
DEHP	DEHP	10
DBP_	DBP	10
DiBP	DiBP	10
BBP	BBP	10
DINP	DINP	10
MCOP	MCOP	10

Table 5 NHANES 2005-2006 Percentiles for Daily Intake (DI; μg/kg/day) values used to calculated Hazard Index. The "DINP" variable uses both MINP and MCOP variable to calculate the DI values. The HI for DINP uses only MCOP to calculated DINP DI as was done in the CHAP report.

Quantiles							
Variable	Label	Percentile		Estimate	Std Error	95% Confidence Limits	
DEHP	DEHP	50%	Median	3.499532	0.286167	2.8895827	4.1094819
	DEHP	90%	D9	18.302700	3.224773	11.4292603	25.1761405
	DEHP	95%		28.508876	9.259956	8.7717464	48.2460049
	DEHP	99%		205.174903	.	.	.
DBP	DBP	50%	Median	0.586159	0.050124	0.4793231	0.6929959
	DBP	90%	D9	2.000158	0.154477	1.6708984	2.3294171
	DBP	95%		2.794720	0.376281	1.9926961	3.5967447
	DBP	99%		5.650076	0.236375	5.1462545	6.1538970
DiBP	DiBP	50%	Median	0.168373	0.010414	0.1461752	0.1905711
	DiBP	90%	D9	0.496134	0.072829	0.3409019	0.6513662
	DiBP	95%		0.892255	0.226408	0.4096782	1.3748325
	DiBP	99%		2.068231	1.717249	-1.5919992	5.7284607
BBP	BBP	50%	Median	0.220460	0.013193	0.1923407	0.2485802
	BBP	90%	D9	0.797963	0.091084	0.6038215	0.9921050
	BBP	95%		1.137141	0.103880	0.9157257	1.3585571
	BBP	99%		2.431174	0.899640	0.5136358	4.3487117
DINP	DINP	50%	Median	0.943162	0.072378	0.7888927	1.0974319
	DINP	90%	D9	4.235127	0.983093	2.1397128	6.3305411
	DINP	95%		8.820979	1.400075	5.8367899	11.8051677
	DINP	99%		27.449664	8.652174	9.0079924	45.8913352
MCOP	MCOP)	50%	Median	0.947751	0.075084	0.7877131	1.1077883
	MCOP	90%	D9	4.838752	1.290844	2.0873831	7.5901211
	MCOP	95%		9.829325	1.686333	6.2349914	13.4236584
	MCOP	99%		25.854974	13.553887	3.0344517	54.7443990

NHANES 2007-2008

NHANES Survey Data 2007-2008 (drop zero daily intake data)
Females Ages 15 to 45

The SURVEYMEANS Procedure

Table 6

Data Summary	
Number of Strata	14
Number of Clusters	17
Number of Observations	57
Number of Observations Used	20
Number of Obs with Nonpositive Weights	37
Sum of Weights	1773266.64

Table 7

Statistics		
Variable	**Label**	**N Miss**
DEHP	DEHP	0
DBP	DBP	0
DiBP	DiBP	0
BBP	BBP	0
DINP	DINP	0
MCOP_	MCOP	0

Table 8 NHANES 2007-2008 Percentiles for Daily Intake (μg/kg/day) values used to calculated Hazard Index. The "DINP" variable uses both MINP and MCOP variable to calculate the DI values. The HI for DINP uses only MCOP to calculated DINP DI as was done in the CHAP report.

Quantiles							
Variable	Label	Percentile		Estimate	Std Error	95% Confidence Limits	
DEHP	DEHP	50%	Median	3.804458	0.795614	1.2724595	6.336456
	DEHP	90%	D9	57.735823	5.991119	38.6694081	76.802239
	DEHP	95%		77.940626	5.991119	58.8742112	97.007042
	DEHP	99%		94.104469	5.991119	75.0380537	113.170884
DBP	DBP	50%	Median	0.661063	0.080148	0.4059950	0.916131
	DBP	90%	D9	1.009151	0.246742	0.2239071	1.794395
	DBP	95%		1.641487	0.112368	1.2838804	1.999093
	DBP	99%		3.204422	0.148472	2.7319168	3.676928
DiBP	DiBP	50%	Median	0.333602	0.079754	0.0797897	0.587415
	DiBP	90%	D9	0.760479	.	.	.
	DiBP	95%		0.766717	.	.	.
	DiBP	99%		0.861848	.	.	.
BBP	BBP	50%	Median	0.467956	0.071962	0.2389418	0.696971
	BBP	90%	D9	1.516870	0.311129	0.5267181	2.507022
	BBP	95%		1.806723	.	.	.
	BBP	99%		2.708253	0.085377	2.4365445	2.979961
DINP	DINP	50%	Median	1.289917	0.136518	0.8554576	1.724377
	DINP	90%	D9	3.047598	0.279234	2.1589524	3.936244
	DINP	95%		5.669005	0.843698	2.9839831	8.354027
	DINP	99%		8.433659	.	.	.
MCOP	MCOP	50%	Median	1.278847	0.181350	0.7017112	1.855983
	MCOP	90%	D9	3.587831	0.369276	2.4126307	4.763032
	MCOP	95%		7.106456	1.114544	3.5594798	10.653432
	MCOP	99%		10.723442	.	.	.

NHANES 2009-2010

NHANES Survey Data 2009-2010 (drop zero daily intake data)
Females Ages 15 to 45

The SURVEYMEANS Procedure

Table 9

Data Summary	
Number of Strata	13
Number of Clusters	20
Number of Observations	68
Number of Observations Used	26
Number of Obs with Nonpositive Weights	42
Sum of Weights	3168615.22

Table 10

Statistics		
Variable	**Label**	**N Miss**
DEHP	DEHP	0
DBP	DBP	0
DiBP	DiBP	0
BBP	BBP	0
DINP	DINP	0
MCOP	MCOP	0

Table 11 NHANES 2009-2010 Percentiles for Daily Intake (μg/kg/day) values used to calculated Hazard Index. The "DINP" variable uses both MINP and MCOP variable to calculate the DI values. The HI for DINP uses only MCOP to calculated DINP DI as was done in the CHAP report.

Quantiles							
Variable	Label	Percentile		Estimate	Std Error	95% Confidence Limits	
DEHP	DEHP	50%	Median	1.616792	0.188248	1.17165570	2.0619286
	DEHP	90%	D9	3.268569	.	.	.
	DEHP	95%		9.232772	.	.	.
	DEHP	99%		9.944895	.	.	.
DBP	DBP	50%	Median	0.511972	0.043360	0.40944198	0.6145018
	DBP	90%	D9	1.003601	.	.	.
	DBP	95%		1.435002	.	.	.
	DBP	99%		5.069528	.	.	.
DiBP	DiBP	50%	Median	0.327473	0.029809	0.25698483	0.3979610
	DiBP	90%	D9	0.501045	.	.	.
	DiBP	95%		0.570342	.	.	.
	DiBP	99%		0.846438	.	.	.
BBP	BBP	50%	Median	0.156970	0.045589	0.04917009	0.2647706
	BBP	90%	D9	0.597319	0.107642	0.34278459	0.8518524
	BBP	95%		0.693235	0.107642	0.43870123	0.9477691
	BBP	99%		1.503062	.	.	.
DINP	DINP	50%	Median	1.755734	0.135952	1.43425940	2.0772089
	DINP	90%	D9	8.443941	0.980536	6.12534261	10.7625391
	DINP	95%		10.492051	0.980536	8.17345319	12.8106497
	DINP	99%		26.375297	.	.	.
MCOP	MCOP	50%	Median	1.994726	0.140837	1.66169848	2.3277540
	MCOP	90%	D9	10.019548	1.507268	6.45542562	13.5836703
	MCOP	95%		13.298412	1.507268	9.73428940	16.8625341
	MCOP	99%		31.867872	.	.	.

NHANES 2011-2012

These data were retrieved before CDC removed the phthalate urinary metabolite data from the website in Aug 2014. The error in the data was with the weighting. These data did have weighting applied therefore the percentiles may be over or under estimated.

NHANES Survey Data 2011-2012 (drop zero daily intake data)
Females Ages 15 to 45

The SURVEYMEANS Procedure

Table 12

Data Summary	
Number of Strata	8
Number of Clusters	12
Number of Observations	56
Number of Observations Used	16
Number of Obs with Nonpositive Weights	40
Sum of Weights	1599267.42

Table 13

Statistics		
Variable	**Label**	**N Miss**
DEHP	DEHP	0
DBP_	DBP	0
DiBP	DiBP	0
BBP	BBP	0
DINP	DINP	0
MCOP	MCOP	0

Table 14 NHANES 2011-2012 Percentiles for Daily Intake (ug/kg/day) values used to calculated Hazard Index. The "DINP" variable uses both MINP and MCOP variable to calculate the DI values. The HI for DINP uses only MCOP to calculated DINP DI as was done in the CHAP report.

Quanti							
Variable	Label	Percentile		Estimate	Std Error	95% Confidence Limits	
DEHP	DEHP	50%	Median	1.904506	0.161792	1.4553001	2.35371212
	DEHP	90%	D9	7.414685	.	.	.
	DEHP	95%		9.310370	.	.	.
	DEHP	99%		12.366212	.	.	.
DBP	DBP	50%	Median	0.319650	0.085142	0.0832567	0.55604312
	DBP	90%	D9	3.926642	0.536088	2.4382223	5.41506100
	DBP	95%		6.020952	0.536088	4.5325329	7.50937157
	DBP	99%		7.696401	0.536088	6.2079813	9.18482003
DiBP	DiBP	50%	Median	0.188726	0.022041	0.1275291	0.24992294
	DiBP	90%	D9	0.497204	.	.	.
	DiBP	95%		0.745200	.	.	.
	DiBP	99%		0.850082	.	.	.
BBP	BBP	50%	Median	0.117243	0.009214	0.0916598	0.14282592
	BBP	90%	D9	1.070864	.	.	.
	BBP	95%		2.229938	.	.	.
	BBP	99%		3.120056	.	.	.
DINP	DINP	50%	Median	4.184664	1.749744	-0.6734044	9.04273158
	DINP	90%	D9	14.522341	.	.	.
	DINP	95%		16.349772	.	.	.
	DINP	99%		18.046377	.	.	.
MCOP	MCOP	50%	Median	5.161099	1.466870	1.0884152	9.23378285
	MCOP)	90%	D9	17.130029	.	.	.
	MCOP	95%		18.967593	.	.	.
	MCOP)	99%		20.784043	.	.	.