July 19, 2024

Comments from Scientists, Academics, and Clinicians on the Draft Risk Evaluations for Di-isodecyl Phthalate (DIDP) and Di-isononyl Phthalate (DINP) Under TSCA

Submitted online via Regulations.gov to docket EPA-HQ-OPPT-2024-0073-0006

These comments are submitted on behalf of the undersigned scientists, academics, and clinicians. We declare that we have no direct or indirect financial or fiduciary interests in the subjects of these comments. The co-signers' institutional affiliations are included for identification purposes only and do not imply institutional endorsement or support.

We appreciate the opportunity to provide written comments on EPA's Draft Risk Evaluation for Di-isodecyl Phthalate, (hereafter referred to as the *DIDP Draft Risk Evaluation*) and the Draft Physical Chemical, Fate, and Hazard Assessments for Di-isononyl Phthalate (hereafter referred to as the *DINP Draft Hazard* Assessment) conducted under the Toxic Substances Control Act ("TSCA"), which requires EPA to evaluate chemical risks based on the "best available science."[1] Both DIDP and DINP are widely used plasticizers, added to a variety of products including building and construction materials, automotive care and fuel products, and consumer products such as adhesives, sealants, paints, coatings, and electrical products.[2] DIDP and DINP are also found in several common household items, such as food packaging materials, nail polishes, fragrances, and pharmaceuticals.[3] Because of their widespread use, both chemicals are considered ubiquitous contaminants that have been detected in most people living in the United States.[4] EPA has identified several non-cancer health hazards of DINP and DIDP exposure, including liver and developmental toxicity.[5]

In both the DIDP Draft Risk Evaluation and the DINP Draft Hazard Assessment, EPA has **failed to incorporate the best available science and makes a number of scientifically-unsupported assumptions that, if adopted, will result in acceptance of serious risks to human health and set a dangerous precedent for future TSCA risk evaluations**. For many conditions of use for DIDP, there are serious inconsistencies between EPA's risk estimates and EPA's conclusions regarding unreasonable risk. EPA also repeatedly downplayed or disregarded the high risks it

[1] 15 USC §2625(h).
[2] U.S. EPA (2024). Draft Risk Evaluation for Diisodecyl Phthalate (DIDP), p. 10.
U.S. EPA (2020). Risk Evaluation for Di-isononyl phthalate (DINP) (1,2-Benzene- dicarboxylic acid, 1,2-diisononyl ester) [Overviews and Factsheets]. https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/risk-evaluation-di-isononyl-phthalate-dinp-12-benzene.
[3] U.S. EPA (2024). Draft Risk Evaluation for Diisodecyl Phthalate (DIDP), p. 10.
U.S. EPA (2020). Risk Evaluation for Di-isononyl phthalate (DINP) (1,2-Benzene- dicarboxylic acid, 1,2-diisononyl ester) [Overviews and Factsheets]. https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/risk-evaluation-di-isononyl-phthalate-dinp-12-benzene.
[4] Zota, A. R., Calafat, A. M., & Woodruff, T. J. (2014). Temporal Trends in Phthalate Exposures: Findings from the National Health and Nutrition Examination Survey, 2001–2010. Environmental Health Perspectives, 122(3), 235–241. https://doi.org/10.1289/ehp.1306681.
[5] U.S. EPA (2024). Draft Non-cancer Human Health Hazard Assessment for Diisononyl Phthalate (DINP), p. 9.
U.S. EPA (2024). Draft Human Health Hazard Assessment for Diisodecyl Phthalate (DIDP), p. 8.

calculated without adequate scientific justification. For example, EPA used mostly central tendency estimates of DIDP exposure and risk for workers and consumers in its unreasonable risk determination, thus disregarding unreasonable risks of non-cancer effects that may be faced by groups with exposures that are greater than median exposure levels. EPA also failed to provide coherent or scientifically-supported rationale for dismissing these risks, an approach that has only been previously employed in the TSCA Draft Risk Evaluation of Formaldehyde. In doing so, EPA continues to set a dangerous precedent that calculated risks can be dismissed or downplayed without scientific support.

In addition, for both draft documents, EPA has failed to adequately consider the scientific evidence and continued to rely on a systematic review methodology that is not consistent with best practices, violating TSCA's "best available science" requirement.[6] For example, both documents improperly excluded all human epidemiological studies from dose-response assessment and relied on systematic review methods that lacked transparency and inappropriately excluded toxicity studies without scientific justification. The National Academies of Sciences, Engineering, and Medicine ("NASEM") has recommended the use of existing systematic review methods and improved approaches for TSCA risk evaluations in 2021, and EPA has still not implemented most of these recommendations.[7] EPA's Science Advisory Committee on Chemicals ("SACC") has also recommended best practices in systematic review to the Agency in multiple reports.[8] EPA should prepare a new TSCA systematic review methodology that is aligned with the best available scientific methods and issue updated draft systematic review protocols for all risk evaluations currently in development, including DIDP and DINP.

Both draft documents also employed a hazard assessment that violates TSCA's "best available science" requirement. While EPA found that developmental and liver toxicity are likely hazards of DIDP and DINP, respectively, it failed to provide quantitative estimates of those non-cancer risk. We applied methods developed by the World Health Organization ("WHO") to quantify the non-cancer risk of developmental toxicity from chronic DIDP exposure, and found that EPA's current approach results in acceptance of exposures producing an upper bound risk of 1-in-100, a risk level *10,000 times higher* than the target risk level that EPA typically applies for protection of carcinogenic risks (1-in-1,000,000). We applied these same methods to quantify the non-cancer risk of liver toxicity from chronic DINP exposure, and found that EPA's current approach results in acceptance of exposures producing an upper bound risk of 1-in-200, a risk level *5,000 times higher* than the typical target risk level.

Another critical concern with the DIDP Draft Risk Evaluation and the DINP Draft Hazard Assessment is EPA's failure to evaluate real world exposures and risks. For example, EPA fails to consider exposures from "non-TSCA" uses for DIDP and DINP, including exposures through food packaging and personal care products. Given that food is the primary route of exposure to both

[6] 15 U.S.C. § 2625(h).
[7] National Academies of Sciences, Engineering, and Medicine (2021). The Use of Systematic Review in EPA's Toxic Substances Control Act Risk Evaluations.
[8] U.S. EPA (2022). Science Advisory Committee on Chemicals Meeting Minutes and Final Report No. 2022-2, p. 71. https://www.regulations.gov/document/EPA-HQ-OPPT-2021-0414-0044.

DINP and DIDP in children and adults,[9] likely as a result of leaching from plastic food packaging materials, EPA will understate the risk to the general population from the TSCA uses of these chemicals if it does not take into account the background exposures from these and other non-TSCA uses. EPA also failed to adequately identify and calculate risks posed to potentially exposed or susceptible subpopulations ("PESS"), as required under TSCA.[10] In the DIDP Draft Risk Evaluation, EPA failed to adequately identify and consider individuals with pre-existing disease, genetic factors, lifestyle factors, geographic factors, or exposures to other chemical and non-chemical stressors that may increase susceptibility to harm from DIDP exposure. A failure to evaluate risk to these groups violates TSCA and results in risk characterization that is not representative of the human population.

Finally, while we support and agree with EPA's decision to conduct a cumulative risk assessment for six phthalates, which includes DINP,[11] without the results of this assessment, EPA cannot make conclusions on hazard or risk in a manner that adequately safeguards human health. In addition, while we agree with EPA's conclusion that DIDP is not antiandrogenic, co- exposures to DIDP and other toxicologically related phthalates may contribute to cumulative risk. As described in the DIDP Draft Risk Evaluation and the 2014 report of the U.S. Consumer Product Safety Commission's Chronic Hazard Advisory Panel on Phthalates and Phthalate Alternatives, DIDP is associated with developmental toxicity.[12] Additional phthalates also are associated with developmental toxicity—including substances that contribute to co-exposures in the human population. If EPA elects not to include DIDP in the proposed cumulative phthalates risk assessment due to the focus on anti-androgenicity, EPA still must address DIDP's potential to contribute to cumulative risk in its individual risk evaluation for DIDP.

Accordingly, EPA must make extensive revisions to both the DIDP Draft Risk Evaluation and the DINP Draft Hazard Assessment to more accurately characterize real-world exposures and risks, including to potentially exposed or susceptible subpopulations. This includes revising its risk determination for DIDP to reflect quantitative non-cancer risk estimates, using high-end exposure and risk estimates, removing the use of any scientifically-unsupported justifications that downplay or disregard risk, and adopting best available scientific methods, like gold-standard systematic review methods that better account for and incorporate the scientific evidence. Furthermore, given EPA's delayed release of the DINP systematic review protocol, EPA should require that an additional public comment period and panel peer review of the DINP hazard assessment documents is conducted following the protocol release.

---

[9] U.S. Consumer Prod. Safety Comm'n, *Report by the Chronic Hazard Advisory Panel on Phthalates and Phthalate Alternatives* 102–03 (2014) at 3, 52–53, 59 (concluding that "food, beverages and drugs via direct ingestion ... constituted the highest [source of] phthalate exposures to all subpopulations").

[10] 15 U.S.C. §§ 2602(12).

[11] US EPA, Draft Proposed Principles of Cumulative Risk Assessment under the Toxic Substances Control Act (Feb. 2023), https://www.epa.gov/system/files/documents/2023-02/Draft%20Principles%20of%20CRA%20under%20TCSA_0.pdf.

[12] U.S. Consumer Prod. Safety Comm'n, *Report by the Chronic Hazard Advisory Panel on Phthalates and Phthalate Alternatives* 102–03 (2014).

Our detailed comments on the DIDP Draft Risk Evaluation and the DINP Draft Hazard Assessment address the following issues:

I. DIDP Draft Risk Evaluation

1. **EPA's non-cancer dose-response assessment for DIDP is not consistent with the best available science.**
    a. **EPA improperly excluded human epidemiology studies from dose-response assessment.**
    b. **EPA failed to apply benchmark dose modeling to derive non-cancer points of departure for risk characterization.**
    c. **EPA's non-cancer margin of exposure (MOE) calculations are unreliable due to EPA's failure to conduct benchmark dose modeling.**
    d. **EPA should apply best available methods to generate quantitative estimates of non-cancer risks for varying levels of exposure to DIDP.**

2. **EPA failed to quantify the cancer risks of DIDP.**
    a. **The best available evidence for DIDP supports a conclusion of "Likely to Be Carcinogenic to Humans"**

3. **EPA did not apply the best available science to identify and evaluate relevant and useful health effects studies for DIDP.**
    a. **EPA did not conduct a comprehensive and up-to-date literature search.**
    b. **EPA inappropriately excluded health effects studies without scientific justification.**
    c. **EPA used deficient inclusion and exclusion criteria for health effects evidence that inappropriately excluded important toxicity endpoints.**
    d. **EPA used multiple strategies to inappropriately exclude PECO-relevant health effects studies.**
    e. **EPA continues to use unclear terminology regarding evidence synthesis and integration.**
    f. **EPA's approach to evidence integration lacks clear procedures and clearly-stated conclusions regarding the hazards of DIDP.**
    g. **EPA released an incomplete draft systematic review protocol for DIDP that was not released in advance of the draft risk evaluation.**
    h. **EPA should prepare a new TSCA systematic review handbook that is aligned with the best available scientific methods and issue updated draft systematic review protocols for all risk evaluations currently in development.**

4. **EPA's occupational and consumer exposure assessments for DIDP are not consistent with the best available science.**
    a. **EPA's determination of unreasonable risk in occupational settings discounts and disregards EPA's own occupational risk estimates for non-cancer effects.**
    b. **EPA inappropriately disregards high-end exposure estimates without justification for its unreasonable risk determinations for workers, ignoring variability in exposures and TSCA's requirement to assess risks to groups with greater exposures.**
    c. **EPA's consumer exposure assessment disregards unreasonable risk to more vulnerable groups, such as infants and toddlers, with little to no scientific justification.**

5. **EPA failed to adequately identify potentially exposed or susceptible subpopulations (PESS), as required by TSCA.**

6. **EPA failed to conduct a background exposure assessment, underestimating risk to potentially exposed or susceptible subpopulations.**

II. DINP Draft Hazard Assessment

1. **EPA's non-cancer dose-response assessment for DINP is not consistent with the best available science.**
    a. **EPA improperly excluded human epidemiology studies from dose-response assessment.**
    b. **EPA failed to apply benchmark dose modeling to derive non-cancer points of departure for risk characterization.**
    c. **EPA should apply best available methods to generate quantitative estimates of non-cancer risks for varying levels of exposure to DINP.**

2. **EPA did not apply the best available science to identify and evaluate relevant and useful health effects studies for DINP.**
    a. **EPA has not released a systematic review protocol for DINP. This means that EPA has employed methods in preparing the DINP hazard assessment that have not been disclosed to the public or to the SACC.**
    b. **EPA did not conduct a comprehensive and up-to-date literature search.**
    c. **EPA relied on assessments conducted by other agencies to exclude studies, without supporting justification.**
    d. **EPA used deficient inclusion and exclusion criteria for health effects evidence that inappropriately excluded important toxicity endpoints.**
    e. **EPA's approach to evidence integration lacks clear procedures and clearly-stated conclusions regarding the hazards of DINP.**

3. **EPA's assessment of DINP carcinogenicity failed to recognize mechanisms in addition to PPARα activation that can contribute to animal liver tumors.**

We appreciate the opportunity to provide public input. Please do not hesitate to contact us with any questions regarding these comments.

Sincerely,

Daniel Axelrad, MPP
Independent Consultant
Washington, DC

Abena BakenRa, MPH
Science Associate, Science and Policy
Program on Reproductive Health and the Environment
University of California, San Francisco

Jessica Trowbridge, PhD, MPH
Associate Research Scientist, Science and Policy
Program on Reproductive Health and the Environment
University of California, San Francisco

Rashmi Joglekar, PhD
Associate Director, Science and Policy
Program on Reproductive Health and the Environment
University of California, San Francisco

Tracey Woodruff, PhD, MPH
Director
Program on Reproductive Health and the Environment
University of California, San Francisco

Ronald H. White, MST
Principal
RHWhite Consulting
Castleton, VA

Ted Schettler, MD, MPH*
Science Director
Science and Environmental Health Network
Bolinas, CA

Rima Habre, ScD, MSc
Associate Professor
Environmental Health and Spatial Sciences
University of Southern California, Population and Public Health Sciences

Jeanne A. Conry, MD, PhD*
President
The Environmental Health Leadership Foundation
Granite Bay, CA

Patrice Sutton, MPH
Research Scientist Collaborator
UCSF Program on Reproductive Health and the Environment
San Francisco, CA

R. Thomas Zoeller, PhD
Emeritus Professor
University of Massachusetts Amherst
Indianapolis, IN

R. Thomas Zoeller, PhD*
Emeritus Professor
The Endocrine Society
Washington D.C.

Mary Zlatnik, MD, MMS
Professor OBGYN
UCSF
San Francisco, CA

Katie Huffling, DNP, RN, CNM, FAAN*
Executive Director
Alliance of Nurses for Healthy Environments
Mount Rainier, MD

Timothy H. Ciesielski, ScD, MD, MPH
Research Scientist
Department of Population and Quantitative Health Sciences
Case Western Reserve University School of Medicine

Laura N. Vandenberg*
Co-chair, Endocrine Society US EDC Task Force
The Endocrine Society
Amherst, MA

Robert M. Gould, MD*
President
San Francisco Bay Physicians for Social Responsibility
San Francisco, CA

Robert M. Gould, MD

Adjunct Assistant Professor
Program on Reproductive Health and the Environment
UCSF Department of Obstetrics, Gynecology and Reproductive Sciences
San Francisco, CA

I. Comments on EPA's DIDP Draft Risk Evaluation

1. EPA's non-cancer dose-response assessment for DIDP is not consistent with the best available science.

a. EPA improperly excluded human epidemiology studies from dose-response assessment.

EPA identified (primarily through public submissions to the docket) more than 25 recent human epidemiology studies of DIDP non-cancer effects that use biomonitoring of urinary metabolites as measures of exposure. EPA improperly excluded all of these studies from dose-response analysis, without any consideration of the strengths and weaknesses of each individual study:

> EPA did not use epidemiology studies quantitatively for dose-response assessment, **primarily due to uncertainty associated with exposure characterization**. Primary sources of uncertainty include uncertainty related to the source of exposure; timing of exposure assessment that may not be reflective of exposure during outcome measurements; co-exposure to mixtures of multiple phthalates that may confound results for the majority of epidemiologic studies, which examine one phthalate and one exposure period at a time such that they are treated as if they occur in isolation; measured urinary metabolites may represent exposure to more than one parent phthalate; and use of spot-urine samples, which due to rapid elimination kinetics may not be representative of average urinary concentrations that are collected over a longer term or calculated using pooled samples.[13]

EPA's blanket exclusion of an entire category of studies is scientifically inappropriate and violates the TSCA requirement to use the "best available science."[14] The preamble to EPA's recent final framework rule for conducting TSCA risk evaluations re-stated EPA's commitment to systematic review:

> EPA believes that integrating appropriate and applicable systematic review methods into the TSCA risk evaluations is critical to meeting the scientific standards as described in TSCA section 26(h) and (i)…. The principles of systematic review are well-established and include "transparent and explicitly documented methods, consistent and critical evaluation of all relevant literature, application of a standardized approach for grading the strength of evidence, and clear and consistent summative language" (Ref. 26). EPA has finalized the requirement to use and document systematic review methods to assess reasonably available information.[15]

EPA's broad exclusion of DIDP epidemiology studies from dose-response analysis is contrary to the framework rule preamble and disregards the structured, consistent systematic review process that is required to evaluate the quality of relevant epidemiological studies according to pre-

[13] U.S. EPA (2024). Draft Human Health Hazard Assessment for Diisodecyl Phthalate (DIDP), p. 13.
[14] 15 U.S.C. § 2625(h).
[15] U.S. EPA (2024). Procedures for Chemical Risk Evaluation Under the Toxic Substances Control Act (TSCA). 89 FR 37028.

specified criteria. EPA has effectively ignored its systematic review process to exclude studies from dose-response assessment with an argument that demonstrates a bias against environmental epidemiology, rather than a thoughtful approach to evidence evaluation that is consistent with best practices in systematic review.

EPA assessed the merits of 26 epidemiology studies of DIDP published from 2018 to 2021 individually, applying a pre-specified set of study quality domains and metrics that closely mirrors the approach used by EPA's IRIS program, which has been favorably reviewed by the National Academies of Sciences, Engineering, and Medicine (NASEM). EPA's overall quality determination was "Medium" or "High" for 25 out of the 26 studies.[16] Each study was individually assessed for its exposure measurement methods (Domain 2) and treatment of potential confounding (Domain 4), and thus the issues that EPA raises in an attempt to disqualify the entire set of epidemiology studies have already been accounted for in a systematic manner using pre-specified procedures. EPA's own overall quality determinations indicate that these studies are suitable for use.

Moreover, EPA's explanation considers only alleged limitations of the DIDP epidemiologic studies as a class, without considering strengths of these studies (e.g., they are conducted in humans rather than laboratory animals, at exposure levels routinely experienced by humans) or mitigating considerations (e.g. regression models that control for co-exposures; implications of exposure misclassification) that apply to the limitations. For example, the use of spot-urine samples is a limitation that is expected to result in some degree of exposure misclassification, but to the extent this occurs, it is likely to result in underestimation of risks. In general, the uncertainties in exposure characterization may result in exposure misclassification that biases dose-response estimates towards the null, but that does not mean the studies are not useful or informative and potentially strong candidates for determination of the POD.

In an attempt to support its decision to disregard epidemiological studies, EPA cites similar decisions made in previous DIDP assessments conducted by other agencies. However, the most recent of these previous assessments considered literature published only up to January 2018, whereas the 26 epidemiology studies assessed for study quality by EPA were all published from 2018-2021, and were therefore not considered in the previous assessments referenced by EPA. EPA does not provide any justification for disregarding its own conclusions regarding these studies when evaluated individually, and by overriding the findings of the systematic review process, EPA therefore violated the TSCA requirement to use the best available science. EPA cannot broadly exclude epidemiologic studies from dose-response assessment in the DIDP Draft Risk Evaluation, and must consider each relevant study on an individual basis as a candidate for POD derivation.

b. **EPA failed to apply benchmark dose modeling to derive non-cancer points of departure for risk characterization.**

[16] U.S. EPA (2024). Draft Risk Evaluation for Diisodecyl Phthalate (DIDP) – Systematic Review Supplemental File: Data Quality Evaluation Information for Human Health Hazard Epidemiology.

EPA violated its own commitment to use EPA guidance in conducting risk evaluations by not applying benchmark dose modeling to derive non-cancer points of departure for risk characterization. EPA has therefore failed to rely on the best available science, and leaves uncertainty regarding whether the most sensitive studies and endpoints were selected for use in estimating risks for DIDP.

For risk characterization of DIDP, EPA proposed to use the no-observed-adverse-effect level (NOAEL) for developmental toxicity (reduced F2 offspring survival) of 38 mg/kg-day (applied dose) from a 2-generation study by Hushka *et al*. for all exposure durations (acute, intermediate and chronic). After application of default allometric scaling, the POD is a human equivalent dose (HED) of 9.0 mg/kg-day.

In Table 6-3 of the DIDP Draft Risk Evaluation, EPA displays 8 NOAEL and LOAEL values for developmental toxicity and liver toxicity that were candidates for the chronic POD.[17] The NOAEL HED values range from 9.0 to 62 mg/kg-day, and the LOAEL HEDs are from 5.2 to 199 mg/kg-day. EPA opted not to use the LOAEL HED of 5.2 mg/kg-day for liver effects from studies by Cho *et al*. because of uncertainties regarding interpretation of the endpoint and higher POD values found when benchmark dose modeling was applied. After setting aside the Cho *et al*. findings, EPA then chose the Hushka *et al*. developmental toxicity NOAEL as the POD because it was more sensitive (i.e., lower) than all other candidate NOAELs and LOAELs.

However, EPA did not conduct benchmark dose (BMD) modeling on Hushka *et al*. or any of the other candidate PODs, except for Cho *et al*. This means that EPA did not apply the best available science to determine the most sensitive endpoint, since it selected the POD without conducting appropriate dose-response analysis. Using a NOAEL as the POD without conducting BMD modeling is not consistent with the best available science, as stated in EPA guidance[18] and reports from the NASEM.[19] By disregarding its own 2012 *Benchmark Dose Technical Guidance* in conducting dose-response analysis for DIDP, EPA has violated its recent final rule *Procedures for Chemical Risk Evaluation Under the Toxic Substances Control Act (TSCA)*, which states:

> EPA will use applicable EPA guidance when conducting risk evaluations, as appropriate and where it represents the best available science.[20]

The Benchmark Dose Technical Guidance document represents the best available science, and it clearly states that NOAELs and LOAELs are significantly limited:

> The NOAEL is actually of little practical utility in describing toxicological dose-response relationships; it does not represent a biological threshold and cannot establish that lower

[17] U.S. EPA (2024). Draft Human Health Hazard Assessment for Diisodecyl Phthalate (DIDP), p. 57.
[18] U.S. EPA (2012). Benchmark Dose Technical Guidance.
[19] NASEM (2017). Application of systematic review methods in an overall strategy for evaluating low-dose toxicity from endocrine active chemicals, p. 158; National Research Council (2009). Science and Decisions: Advancing Risk Assessment, p. 129.
[20] 40 CFR § 702.37

exposure levels are necessarily without risk. Specific limitations of the NOAEL/LOAEL approach are well known and have been discussed extensively (Crump 1984; Gaylor 1983; Kimmel and Gaylor 1988; Leisenring and Ryan 1992; U.S. EPA 1995a):

- The NOAEL/LOAEL is highly dependent on sample size. The ability of a bioassay to distinguish a treatment response from a control response decreases as sample size decreases, so the NOAEL for a compound (and thus the POD, when based on a NOAEL) will tend to be higher in studies with smaller numbers of animals per dose group.
- More generally, the NOAEL/LOAEL approach does not account for the variability and uncertainty in the experimental results that are due to characteristics of the study design such as dose selection, dose spacing, and sample size.
- NOAELs/LOAELs do not correspond to consistent response levels for comparisons across studies/chemicals/endpoints, and the observed response level at the NOAEL or LOAEL is not considered in the derivation of RfDs/RfCs.
- Other dose-response information from the experiment, such as the shape of the dose-response curve (e.g., how steep or shallow the slope is at the BMD, providing some indication of how near the POD might be to an inferred threshold), is not taken into account…
- While the NOAEL has typically been interpreted as a threshold (no-effect level), simulation studies (e.g., Leisenring and Ryan 1992; study designs involving 10, 20, or 50 replicates per dose group) and re-analyses of developmental toxicity bioassay data (Gaylor 1992; Allen et al. 1994a; studies involving approximately 20 litters per dose group) have demonstrated that the rate of response above control at doses fitting the criteria for NOAELs, for a range of study designs, is about 5–20% on average, not 0%.[21]

The Benchmark Dose Technical Guidance further states that use of a BMD/BMDL as a POD is preferred and a NOAEL or LOAEL should be considered as a POD only if BMD modeling is conducted and is unable to produce a BMD estimate, and requires justification:

> Because of the limitations of the NOAEL/LOAEL approach discussed earlier, the BMD approach is preferred to the NOAEL/LOAEL approach… there are some instances in which reliable BMDs cannot be estimated and the NOAEL/LOAEL approach might be warranted…In such cases, the NOAEL/LOAEL approach might be used, while recognizing its limitations and the limitations of the dataset.[22]

> Resorting to the NOAEL/LOAEL approach does not resolve a data set's inherent limitations, but it conveys that there are limitations with the data set.[23]

[21] U.S. EPA (2012). Benchmark Dose Technical Guidance, p. 4.
[22] U.S. EPA (2012). Benchmark Dose Technical Guidance, p. 6.
[23] U.S. EPA (2012). Benchmark Dose Technical Guidance, p. 12.

> At times, modeling will not yield useful results and the NOAEL/LOAEL approach might be considered, although the data gaps and inherent limitations of that approach should be acknowledged.[24]
>
> In some cases, modeling attempts may not yield useful results. When this occurs and the most biologically relevant effect is from a study considered adequate but not amenable to modeling, the NOAEL (or LOAEL) could be used as the POD. The modeling issues that arose should be discussed in the assessment, along with the impacts of any related data limitations on the results from the alternate NOAEL/LOAEL approach.[25]

In addition, the Benchmark Dose Technical Guidance clearly states that identification of the most sensitive endpoint cannot be based on comparisons of NOAELs, and all candidate values should be evaluated based on BMD modeling:

> The apparent relative sensitivities of endpoints based on NOAELs/LOAELs may not correspond to the same relative sensitivities based on BMDs or BMDLs after BMD modeling; therefore, relative sensitivities of endpoints cannot necessarily be judged a priori. For example, differences in slope (at the BMR) among endpoints could affect the relative values of the BMDLs. Selected endpoints from different studies that have the potential to be used in the determination of a POD(s) should all be modeled.[26]

EPA cited the Benchmark Dose Technical Guidance in a previous TSCA risk evaluation to describe the preference for a BMD over a NOAEL:

> As outlined in EPA guidance, the BMD approach overcomes many of the limitations inherently associated with the NOAEL/LOAEL approach, and thus is the preferred method for establishing a POD for use in risk assessment.[27]

EPA's 2022 guidance for conducting chemical hazard assessments for the Integrated Risk Information System (IRIS) reinforces these key points:

> As discussed in detail in Section 1.2 of EPA's Benchmark Dose Technical Guidance (U.S. EPA, 2012b), dose-response modeling (i.e., benchmark dose modeling) is the preferred approach for deriving points of departures given several limitations in the no-observed adverse-effect level/ lowest-observed-adverse-effect level (NOAEL/LOAEL) approach.[28]
>
> Basis of the POD: A modeled BMDL is preferred over a NOAEL, which is in turn preferred over a LOAEL.[29]

---

[24] U.S. EPA (2012). Benchmark Dose Technical Guidance, p. 30.
[25] U.S. EPA (2012). Benchmark Dose Technical Guidance, p. 40.
[26] U.S. EPA (2012). Benchmark Dose Technical Guidance, p. 15.
[27] U.S.EPA (2020). Risk Evaluation for n-Methylpyrrolidone (2-Pyrrolidinone, 1-Methyl-) (NMP), p. 262.
[28] U.S. EPA (2022). ORD Staff Handbook for Developing IRIS Assessments, p. 8-1.
[29] U.S. EPA (2022). ORD Staff Handbook for Developing IRIS Assessments, p. 8-18.

Reports from the NASEM also state the advantages of BMD modeling. The NASEM report on low-dose toxicity of endocrine active chemicals discusses the deficiencies of the NOAEL/LOAEL approach for risk estimation:

> The use of LOAELs and NOAELs is less than ideal because they depend highly on individual study-design characteristics; therefore, apparent differences among studies might be explained by design differences, such as sample size or dose spacing, rather than true inconsistency.[30]

In the 2009 report *Science and Decisions*, the National Academies highlighted the adoption of the BMD approach as an important improvement in risk assessment methodology:

> Another refinement in dose-response assessment has been the derivation of the RfD or low-dose cancer risk from a POD that is calculated using BMD methodology (EPA 2000a). In noncancer risk assessment, this approach has the advantage of making better use of the dose-response evidence available from bioassays than do calculations based on NOAELs. It also provides additional quantitative insight into the risk presented in the bioassay at the POD because for quantal end points the POD is defined in terms of a given risk for the animals in the study.[31]

EPA selected the NOAEL for developmental toxicity of 9.0 mg/kg-day (HED) from Hushka *et al.* as the chronic exposure POD because it has the lowest NOAEL (i.e., it is the "most sensitive" outcome) among the several candidate values, as shown in Table 6-3.[32] EPA correctly identifies this outcome as "clearly adverse."[33,34] However, EPA has not conducted the BMD modeling necessary to determine if the other endpoints are more sensitive. A BMDL is frequently lower than the NOAEL for the same endpoint, and frequently much lower than the LOAEL for the same endpoint. Without BMD modeling, EPA is unable to make a scientific determination of whether this Hushka *et al.* endpoint is more or less sensitive than the liver toxicity endpoint from Hazelton Labs (NOAEL HED = 9.3 mg/kg-day), the liver toxicity endpoint from Hushka *et al*. (P2 males, NOAEL HED = 28 mg/kg-day), developmental toxicity endpoints from Hushka *et al.* Study A (*LOAEL* HED = 32 mg/kg-day), or other candidate endpoints. The scientifically appropriate method for selecting the POD based on the most sensitive endpoint would be to first estimate a BMDL for each endpoint, and then select the lowest value.

As part of conducting BMD modeling, it is critical that EPA select an appropriate benchmark response (BMR) for each endpoint. For severe effects, such as reduced neonatal survival, EPA should apply a BMR of no more than 1%, consistent with past practice: for example, the TSCA risk evaluation for n-methylpyrrolidone (NMP) applied a BMR of 1% for post-implantation

[30] NASEM (2017). Application of systematic review methods in an overall strategy for evaluating low-dose toxicity from endocrine active chemicals, p. 158.
[31] National Research Council (2009). Science and Decisions: Advancing Risk Assessment, p. 129.
[32] U.S. EPA (2024). Draft Human Health Hazard Assessment for Diisodecyl Phthalate (DIDP), p. 57.
[33] U.S. EPA (2024). Draft Risk Evaluation for Diisodecyl Phthalate (DIDP), p. 110.
[34] U.S. EPA (2024). Draft Human Health Hazard Assessment for Diisodecyl Phthalate (DIDP), p. 9.

losses and stillbirths,[35] and the TSCA risk evaluation for trichloroethylene (TCE) applied a BMR of 1% for fetal heart malformations.[36]

By disregarding existing EPA guidance and NASEM recommendations that state BMD modeling is the most scientifically appropriate approach for determining the POD, EPA has violated TSCA section 26(h), which direct that the Agency:

> Shall use scientific information, technical procedures, measures, methods, protocols, methodologies, or models, employed in a manner consistent with the best available science.[37]

EPA's recently promulgated revisions to the framework rule for TSCA risk evaluations states that:

> EPA will document that the risk evaluation is consistent with the best available science.[38]

EPA cannot ensure that the final DIDP risk evaluation meets this requirement unless it has implemented BMD modeling in the process of selecting a POD.

c. EPA's non-cancer margin of exposure (MOE) calculations are unreliable due to EPA's failure to conduct benchmark dose modeling.

To inform its determination of unreasonable risks of non-cancer effects from chronic exposure, EPA calculated a margin of exposure (MOE) for each DIDP condition of use (COU) using the POD (HED) of 9.0 mg/kg-day. The MOE is calculated as:

> Margin of Exposure = Non-cancer point of departure / Human exposure.

As discussed below, the MOE approach is a scientifically deficient method for characterizing risk and is inconsistent with amended TSCA's requirements to use the "best available science"[39] and to ensure protection of "potentially exposed and susceptible subpopulations" ("PESS").[40]

In the DIDP Draft Risk Evaluation, the many shortcomings of EPA's MOE approach are exacerbated by EPA's failure to conduct dose-response modeling. EPA's calculated MOEs for DIDP are in question because of EPA's use of a NOAEL as the POD. Application of BMD modeling could result in a POD that is significantly lower than the NOAEL, which in turn would significantly reduce the calculated MOEs. COUs that currently have calculated MOEs up to 100 or even greater could conceivably be reduced to below EPA's benchmark MOE of 30 when recalculated with an appropriate POD, and should be provisionally considered contributors to unreasonable risk until EPA has conducted BMD modeling of multiple non-cancer endpoints

[35] U.S. EPA (2020). Risk Evaluation for n-Methylpyrrolidone (2-Pyrrolidinone, 1-Methyl-) (NMP).
[36] U.S. EPA (2020). Risk Evaluation for Trichloroethylene.
[37] 15 U.S.C. § 2625(h).
[38] U.S. EPA (2024). Procedures for Chemical Risk Evaluation Under the Toxic Substances Control Act (TSCA). 89 FR 37028, May 3, 2024, § 702.37(a)(2).
[39] 15 U.S.C. §2625 (h).
[40] 15 U.S.C. §2602 (12).

and, preferably, conducted a probabilistic dose-response analysis, as described below, to replace the MOE approach.

In addition, the Draft Occupational Exposure Value Calculations in Appendix F are similarly not scientifically defensible due to the failure to conduct BMD modeling in selecting a POD.[41] At a minimum, the draft occupational exposure value must be recalculated after conducting BMD modeling for multiple candidate endpoints and selection of a POD based on the BMDL values.

d. **EPA should apply best available methods to generate quantitative estimates of non-cancer risks for varying levels of exposure to DIDP.**

As discussed above, the draft DIDP risk evaluation continues EPA's practice of relying on the scientifically deficient MOE approach for non-cancer dose-response analysis and risk characterization in TSCA risk evaluations. The MOE approach does not provide a quantitative estimate of risk and is inconsistent with TSCA's requirements to use the "best available science"[42] and to ensure protection of "potentially exposed and susceptible subpopulations" ("PESS").[43]

Use of the MOE, which relies on a POD with no extrapolation to lower doses, is a simplistic approach that only compares the POD to the exposure level and judges whether this ratio is interpreted as a human health risk of concern or if "risk is not considered to be of concern and mitigation is not needed."[44] The MOE does not estimate the proportion of the exposed population projected to experience a specified health endpoint or the number of individuals affected, and it perpetuates the scientifically flawed notion that a "safe" or "no risk" level of chemical exposure can be identified for a diverse exposed population.[45,46]

The National Academies[47] and the World Health Organization[48] (WHO) have outlined more robust methods for risk estimation that more accurately account for variability and vulnerability

[41] U.S. EPA (2024). Draft Risk Evaluation for Diisodecyl Phthalate (DIDP), p. 221.
[42] 15 U.S.C. §2625 (h).
[43] 15 U.S.C. §2602 (12).
[44] U.S. EPA (2024). Draft Risk Evaluation for Diisodecyl Phthalate (DIDP), p. 113.
[45] Woodruff, T. J., Rayasam, S. D. G., Axelrad, D. A., Koman, P. D., Chartres, N., et al.. (2023). A science-based agenda for health-protective chemical assessments and decisions: overview and consensus statement. Environ Health, 21(Suppl 1), 132. https://doi.org/10.1186/s12940-022-00930-3.
[46] McGartland, A., Revesz, R., Axelrad, D. A., Dockins, C., Sutton, P., Woodruff, T. J. (2017). Estimating the health benefits of environmental regulations. Science, 357(6350), 457-458. https://doi.org/10.1126/science.aam8204.
[47] National Research Council (2009). Science and Decisions: Advancing Risk Assessment, Chapter 5.
[48] WHO. (2017). Guidance document on evaluating and expressing uncertainty in hazard characterization. Harmonization project document 11, 2nd edition. https://www.who.int/publications/i/item/9789241513548.

across the human population and have been demonstrated in published case studies.[49,50,51,52] We applied the WHO methodology to the DIDP endpoint of reduced offspring survival to estimate risk-specific doses for several levels of incidence (e.g. 1%, 0.1%, etc.) and at doses relevant to the DIDP Draft Risk Evaluation. Because EPA has not estimated BMDLs for the chronic effects of DIDP, we use the NOAEL from the critical study identified by EPA as the starting point for this analysis.

Our analysis (see Technical Appendix A for details) found that:

1. 1.0 mg/kg-day is the lower bound (95% confidence) chronic human dose at which clearly adverse developmental effects are expected in 10% of the exposed population;
2. 0.27 mg/kg-day is the lower bound (95% confidence) chronic human dose at which clearly adverse developmental effects are expected in 1% of the exposed population;
3. 0.1 mg/kg-day is the lower bound (95% confidence) chronic human dose at which clearly adverse developmental effects are expected in 0.1% of the exposed population;
4. 0.04 mg/kg-day is the lower bound (95% confidence) chronic human dose at which clearly adverse developmental effects are expected in 0.01% (1-in-10,000) of the exposed population;
5. 0.02 mg/kg-day is the lower bound (95% confidence) chronic human dose at which clearly adverse developmental effects are expected in 0.001% (1-in-100,000) of the exposed population.

The implications of these risk values can be understood by comparison with the exposure levels considered by EPA to represent negligible risk. EPA's assessment uses a POD of 9.0 mg/kg-day and a benchmark MOE of 30, meaning that EPA concludes "risk is not considered to be of concern and mitigation is not needed"[53] for any exposure below 0.30 mg/kg-day (9.0 mg/kg-day / 30 = 0.30 mg/kg-day). Our analysis finds that an exposure of 0.30 mg/kg-day exceeds the lower-bound dose for the 1% (1-in-100) risk level. This risk far exceeds EPA's usual target range of protection for carcinogenic risks of 1-in-10,000 to 1-in-1,000,000.[54]

The risk values obtained from application of the WHO framework also indicate that many workers are at high risk for adverse non-cancer effects:

---

[49] Chiu WA, Axelrad DA, Dalaijamts C, Dockins C, Shao K, Shapiro AJ, Paoli G. Beyond the RfD: broad application of a probabilistic approach to improve chemical dose-response assessment for non-cancer effects. Environmental Health Perspectives, 2018 June;126(6):067009. doi:10.1289/EHP3368.

[50] Nielsen, G. H., Heiger-Bernays, W. J., Levy, J. I., White, R. F., Axelrad, D. A., Lam, J., Chartres, N., Abrahamsson, D. P., Rayasam, S. D. G., Shaffer, R. M., Zeise, L., Woodruff, T. J., Ginsberg, G. L. (2023). Application of probabilistic methods to address variability and uncertainty in estimating risks for non-cancer health effects. Environ Health, 21(Suppl 1), 129. https://doi.org/10.1186/s12940-022-00918-z.

[51] Blessinger, T., Davis, A., Chiu, W. A., Stanek, J., Woodall, G. M., Gift, J., Thayer, K. A., Bussard, D. (2020). Application of a unified probabilistic framework to the dose-response assessment of acrolein. Environ Int, 143,105953. https://doi.org/10.1016/j.envint.2020.105953.

[52] Ginsberg, G. L. (2012). Cadmium risk assessment in relation to background risk of chronic kidney disease. J Toxicol Environ Health A, 75(7),374-390.

[53] U.S. EPA (2024). Draft Risk Evaluation for Diisodecyl Phthalate (DIDP), p. 113.

[54] U.S. EPA (2024). Unreasonable Risk Determination of the Draft Risk Evaluation for Formaldehyde, p. 13.

- High-end exposure estimates for 3 occupational exposure scenarios are greater than or equal to 0.27 mg/kg, the lower-bound dose estimate for 1% (1-in-100) risk: application of adhesives and sealants, application of paints and coatings, and use of penetrants and inspection fluids;
- High-end exposure estimates for an additional 5 occupational exposure scenarios are greater than or equal to 0.1 mg/kg, the lower-bound dose estimate for 0.1% (1-in-1,000) risk: PVC plastics compounding, PVC plastics converting, non-PVC material converting, recycling, and disposal;
- Central-tendency exposure estimates for 5 occupational exposure scenarios are greater than or equal to 0.04 mg/kg, the lower-bound dose estimate for 0.01% (1-in-10,000) risk: application of adhesives and sealants, application of paints and coatings, use of penetrants and inspection fluids, PVC plastics compounding, and non-PVC material converting.

EPA should apply the WHO framework to the reduced offspring survival endpoint of DIDP using a BMD_{01} instead of the NOAEL as the starting point, and should also apply the framework to other non-cancer endpoints of DIDP with BMD-derived PODs for comparison.

2. EPA failed to quantify the cancer risks of DIDP.

a. The best available evidence for DIDP supports a conclusion of "Likely to Be Carcinogenic to Humans"

EPA's review of the evidence for carcinogenicity of DIDP finds that leukemias were significantly elevated in DIDP-exposed male and female rats, and hepatocellular adenomas were significantly elevated in DIDP-exposed male mice. EPA concluded that the evidence supports a conclusion of "Suggestive Evidence of Carcinogenic Potential."[55]

Among the types of evidence cited in EPA's 2005 *Guidelines for Carcinogen Risk Assessment* that support a conclusion of "Likely to Be Carcinogenic to Humans" are positive findings in multiple species and sexes:

> an agent that has tested positive in animal experiments in more than one species, sex, strain, site, or exposure route, with or without evidence of carcinogenicity in humans.[56]

With positive findings in male rats, female rats, and male mice, DIDP meets the requirements for a designation of "Likely" rather than "Suggestive" evidence. EPA should revise the conclusion, conduct dose-response analysis to estimate the cancer potency of DIDP, estimate cancer risks to workers, consumers and the general population, and identify any conditions of use with cancer risks greater than 1-in-1,000,000 to be contributors to unreasonable risk.

[55] U.S. EPA (2024). Draft Human Health Hazard Assessment for Diisodecyl Phthalate (DIDP), p. 48.
[56] U.S. EPA (2005). Guidelines for Carcinogen Risk Assessment, p. 2-55.

3. EPA did not apply the best available science to identify and evaluate relevant and useful health effects studies for DIDP.

a. EPA did not conduct a comprehensive and up-to-date literature search.

The need for transparent, consistent and comprehensive approaches to identifying health effects literature has been a key driver for increased adoption of systematic review methods in environmental health assessments over the past 15 years.[57,58,59] EPA's assessment of DIDP is a concerning step backwards in this area, as the approach to identifying evidence is not clear, consistent or comprehensive. Based on the inconsistent procedures applied, it is unlikely that EPA would have identified and included all relevant health effects studies. This indicates critical deficiencies in the EPA systematic review protocol and the DIDP Draft Risk Evaluation.

For the DIDP Draft Risk Evalution, EPA relied on non-EPA assessments of DIDP completed in 2018 or earlier, and a literature search that was conducted in 2019 and has not been updated since.

For identifying epidemiological studies, EPA described its procedures as follows:

> To identify and integrate human epidemiologic data into the draft DIDP Risk Evaluation, EPA first reviewed existing assessments of DIDP conducted by regulatory and authoritative agencies…**most** of these assessments have been subjected to external peer-review and/**or** public comment periods, but **have not employed formal systematic review protocols**.[60] (emphasis added)

> Next, EPA sought to identify new population, exposure, comparator, and outcome (PECO)-relevant literature published since the most recent existing assessment of DIDP. PECO-relevant literature published since the most recent existing assessment(s) of DIDP was identified by applying a literature inclusion cutoff date from existing assessments of DIDP. For DIDP, EPA used the applied cutoff date based on existing assessments of epidemiologic studies of phthalates by Health Canada (2018a, b), which included literature up to January 2018...New PECO-relevant literature published between 2018 to 2019 that was identified through the literature search conducted by EPA in 2019, as well as references published between 2018 to 2023 that were submitted with public comments to the DIDP Docket…were evaluated for data quality.[61]

---

[57] National Research Council (2011). Review of the Environmental Protection Agency's Draft IRIS Assessment of Formaldehyde.
[58] Woodruff TJ, Sutton P; Navigation Guide Work Group. An evidence-based medicine methodology to bridge the gap between clinical and environmental health sciences. Health Affairs 2011 May;30(5):931-7. doi: 10.1377/hlthaff.2010.1219.
[59] Rooney AA, Boyles AL, Wolfe MS, Bucher JR, Thayer KA. 2014. Systematic review and evidence integration for literature-based environmental health science assessments. Environ Health Perspect 122:711–718.
[60] U.S. EPA (2024). Draft Human Health Hazard Assessment for Diisodecyl Phthalate (DIDP), p. 11.
[61] U.S. EPA (2024). Draft Human Health Hazard Assessment for Diisodecyl Phthalate (DIDP), p. 12.

EPA therefore conducted a comprehensive literature search only for studies published in a time period of less than 2 years. As a result, the set of epidemiology studies consists of three inconsistent subsets:

- Studies published prior to January 2018 – are included in EPA's assessment only if they were included in the assessments conducted by other agencies. The assessments used by EPA to identify studies were not necessarily peer-reviewed and were not systematic reviews. EPA did not assess the quality of the studies identified by these other assessments. EPA did not consider any studies published before 2018 if they were not discovered by or not included in previous assessments for any reason.
- Studies published from January 2018 – September 2019 – EPA conducted its own search of the literature and applied its own inclusion/exclusion criteria, followed by a "further filtering" process (discussed below) and study quality evaluation procedures.
- Studies published from September 2019 to May 2024 – are included in EPA's assessment only if they were submitted to the EPA docket.

Thus, only those epidemiology studies published in a 21-month span were identified and evaluated through a comprehensive process following an EPA protocol. For earlier studies (before 2018), EPA relied entirely on the Health Canada and other agency assessments and did not apply its own search, inclusion/exclusion and study evaluation procedures. For later studies (after September 2019), EPA did not conduct a search but included only those studies that were submitted by the public to EPA. This is not a clear, comprehensive or consistent approach to identifying the epidemiological evidence relevant to assessing the health effects of DIDP. A further concern is that these inconsistent procedures for identifying epidemiological evidence were ultimately relevant only to the identification of DIDP hazards, since EPA subsequently excluded **all** epidemiological studies from consideration for dose-response assessment, without consideration of the merits of individual studies (see Section 1a. above).

For identifying toxicology studies, EPA applied a similar process:

> EPA first reviewed existing assessments of DIDP conducted by various regulatory and authoritative agencies…The purpose of this review was to identify sensitive and human relevant hazard outcomes associated with exposure to DIDP, and identify key studies used to establish PODs for extrapolating human risk.[62]
>
> EPA…identified PECO-relevant literature published since two recent and comprehensive existing assessments of DIDP by applying a literature inclusion cutoff date from these assessments. For DIDP, assessments by Health Canada (EC/HC, 2015) and Australia NICNAS (NICNAS, 2015) included literature up to August 2014 and July 2014, respectively… assessments by both Health Canada and NICNAS were subject to public comment periods and the assessment by Health Canada was subject to external peer-review…Therefore, EPA considered literature published between 2014 to 2019 further…

[62] U.S. EPA (2024). Draft Human Health Hazard Assessment for Diisodecyl Phthalate (DIDP), p. 13.

EPA reviewed new studies published between 2014 and 2019 and extracted key study information.[63]

EPA therefore conducted a comprehensive literature search only for studies published in a 5-year span. As a result, DIDP toxicology studies are divided into three inconsistent subsets:

- Studies published up to mid-2014 – included only if they were included in previous assessments conducted by Canada and Australia and others. However, according to Figure 4-6 of the DIDP systematic review protocol, EPA included only five studies out of 31 studies that were included in the Canada and Australia assessments. No explanation is provided for the exclusion of the other 26 studies. Additionally, EPA did not consider any studies published before mid-2014 if they were not discovered by or not included in the previous assessments for any reason.
- PECO-relevant studies published from mid-2014 to September 2019 – included in the draft risk evaluation only if they satisfy a "further filtering" process, discussed below.
- Studies published after September 2019 – were not considered at all.

Thus, only those toxicology studies published in a 5-year span were identified and evaluated through a comprehensive process following an EPA protocol. For earlier toxicology studies (before mid-2014), EPA relied entirely on assessments by other agencies and did not apply its own search, inclusion/exclusion and study evaluation procedures. Toxicology studies published after September 2019 were not included at all. This is not a clear, comprehensive or consistent approach to identifying the toxicology evidence relevant to assessing the health effects of DIDP.

For both epidemiology and toxicology, studies were treated differently based only on their date of publication. In addition, the procedures for epidemiology differed significantly from those for toxicology; for example, some post-2019 epidemiology studies were included (but not necessarily all relevant studies, since a search was not conducted), whereas no post-2019 toxicology studies were included. Any toxicological findings on DIDP published in the past 5 years were simply not considered by EPA, which is not consistent with the best available science; recent guidance on conducting systematic reviews in environmental health recommends that literature searches should be updated no more than 12 months before publication of a review.[64] Collectively, EPA's practices run a high risk of failing to include all relevant health effects studies and/or treating relevant studies differently in the DIDP Draft Risk Evaluation.

b. EPA inappropriately excluded health effects studies without scientific justification.

EPA reviewed DIDP health effects assessments conducted by Canada, Australia, multi-lateral European agencies, the U.S. CPSC and the U.S. NTP as part of conducting the DIDP Draft Risk Evaluation. Epidemiology studies published before 2019 and toxicology studies published

[63] U.S. EPA (2024). Draft Human Health Hazard Assessment for Diisodecyl Phthalate (DIDP), pp. 13-14.
[64] Whaley, et al. Recommendations for the conduct of systematic reviews in toxicology and environmental health research (COSTER). Environment International 143 (2020), 105926. https://doi.org/10.1016/j.envint.2020.105926.

before mid-2014 were included in the TSCA risk evaluation only if they were included in these previous assessments. EPA did not assess the quality of studies that were included in the TSCA risk evaluation based on the other agencies' assessments. Studies that were not identified in searches conducted in the previous assessments and studies that were excluded from the previous assessments for any reason were not considered at all by EPA.

In principle, the use of previous assessments can be a useful part of conducting a TSCA risk evaluation, but the previous assessments must be carefully evaluated against a pre-specified set of criteria to determine whether they are of sufficient quality, and the resulting risk evaluation must still employee procedures that are transparent, comprehensive, consistent and unbiased, and must meet the TSCA section 26(h) scientific standards which direct that the Agency:

> Shall use scientific information, technical procedures, measures, methods, protocols, methodologies, or models, employed in a manner consistent with the best available science.[65]

However, EPA notes that the previous assessments it used were not systematic reviews, and not all were peer reviewed. EPA also does not provide adequate justification for its use of previous DIDP assessments to substitute for conducting its own comprehensive systematic review to identify and evaluate health effects evidence.

The 2023 NASEM report *Building Confidence in New Evidence Streams for Human Health Risk Assessment* demonstrates an appropriate process for evaluating the quality of previous assessments. After conducting a comprehensive search for prior reviews satisfying a pre-specified PECO (population, exposure, comparator, outcome) statement, the NASEM applied AMSTAR 2 (A MeaSurement Tool to Assess systematic Reviews) to assess the methodological quality of each relevant review.[66] AMSTAR 2 was also applied by the NASEM in multiple prior reports on environmental health assessment.[67,68,69] In order to establish that it is appropriate to use previous assessments as part of a TSCA risk evaluation, EPA must apply this type of process to determine whether the previous assessments are consistent with the best available science, as required by TSCA.

c. EPA used deficient inclusion and exclusion criteria for health effects evidence that inappropriately excluded important toxicity endpoints.

EPA's PECO statement for DIDP health effects studies is not included in the draft DIDP systematic review protocol. The PECO statement for DIDP is only available only in the broader 2021 TSCA Draft Systematic Review Protocol, which EPA has never revised to address public comments and more than 200 SACC recommendations. PECO statements play a critical role in conducting a systematic review as they provide criteria for screening the literature search results

[65] 15 U.S.C. § 2625(h).
[66] NASEM 2023). Building Confidence in New Evidence Streams for Human Health Risk Assessment: Lessons Learned from Laboratory Mammalian Toxicity Tests.
[67] NASEM (2019). Review of DOD's Approach to Deriving an Occupational Exposure Level for Trichloroethylene.
[68] NASEM (2021). The Use of Systematic Review in EPA's Toxic Substances Control Act Risk Evaluations.
[69] NASEM (2022). Guidance on PFAS Exposure, Testing, and Clinical Follow-Up.

to identify which studies are relevant (included in the risk evaluation) and not relevant (excluded from further consideration). The PECO statement for DIDP is deficient and excludes a broad range of important toxicity outcomes from consideration in the DIDP Draft Risk Evaluation.

The outcome component of the PECO statement for DIDP health effects evidence provides the following criteria for inclusion and exclusion of studies:

> **Human:** All health outcomes (cancer and non-cancer) at the organ level or higher.
> **Animal and Plants:** All apical biological effects (effects measured at the organ level or higher)
> and bioaccumulation from laboratory studies with concurrently measured media and/or tissue concentrations. Apical endpoints include but are not limited to reproduction, survival, and growth.
> **Screener note:**
> - Measurable biological effects relevant for humans, animals and plants may include but are not limited to: mortality, behavioral, population, physiological, growth, reproduction, systemic, point of contact (irritation and sensitization) effects.
> - Effects measured at the cellular level of biological organization and below are to be tagged as supplemental, mechanistic.[70] (emphasis added)

By limiting the relevant human and animal studies to those with "apical" effects, or those with effects at the "organ level or higher," EPA appears to be excluding studies of important biochemical markers and other outcomes at the cellular level that are strong indicators of hazards and that have commonly been used as critical effects in previous EPA hazard assessments, including TSCA risk evaluations (see examples below).

EPA's PECO statement also provides very limited guidance for screeners on what effects are to be considered "apical" or "organ-level." The PECO says: "Apical endpoints include but are not limited to reproduction, survival, and growth" and "Measurable biological effects relevant for humans, animals and plants may include but are not limited to: mortality, behavioral, population, physiological, growth, reproduction, systemic, point of contact (irritation and sensitization) effects."[71] The 2021 TSCA Draft Systematic Review Protocol provides no further guidance on which outcomes are to be considered apical or organ-level, and which outcomes are to be considered cellular-level.

The NASEM has defined an apical end point as "An observable outcome in a whole organism, such as a clinical sign or pathologic state, that is indicative of a disease state that can result from exposure to a toxicant,"[72] and identified "tumors, birth defects, and neurologic impairments"[73] as

[70] U.S. EPA (2021). Draft Systematic Review Protocol Supporting TSCA Risk Evaluations for Chemical Substances, Table_Apx H-47.

[71] U.S. EPA (2021). Draft Systematic Review Protocol Supporting TSCA Risk Evaluations for Chemical Substances, Table_Apx H-47.

[72] National Research Council (2007). Toxicity Testing in the 21st Century: A Vision and a Strategy, p. 38.

[73] National Research Council (2007). Toxicity Testing in the 21st Century: A Vision and a Strategy, p. 177.

examples. No biochemical measures or early biological changes were mentioned among the examples.

The definition of an apical effect appears to be narrower than the definition of an adverse effect provided by the EPA IRIS program: "a biochemical change, functional impairment, or pathologic lesion that affects the performance of the whole organism, or reduces an organism's ability to response to an additional environmental challenge."[74] The definition of adverse effect includes, for example, "a biochemical change;" such effects appear to be excluded from the DIDP Draft Risk Evaluation as they would likely be considered cellular-level effects rather than organ-level or apical effects

Biochemical and/or cellular-level outcomes have been identified as critical effects in numerous past EPA hazard assessments, including some of the completed TSCA risk evaluations. Examples of these outcomes and past assessments include:

- reduced male fetal testosterone or adult male testosterone levels (2018 and 2019 IRIS staff published systematic reviews of health effects of phthalates, 2023 draft approach to cumulative risk assessment of phthalates under TSCA).[75,76,77]
- reduced thyroid hormone levels (2020 TSCA risk evaluation of HBCD; 2021 toxicity assessment of PFBS) [78,79]
- decreased erythrocyte counts and hemoglobin (2020 TSCA risk evaluation of perchloroethylene)[80]
- measures of immune function, such as increases in immunoglobulin E, lymphocytes, natural killer cells, and interlukin-4 levels (2020 TSCA risk evaluation of perchloroethylene)[81]
- decreased sperm quality or concentration (2020 TSCA risk evaluations of trichloroethylene and perchloroethylene; 2018 and 2019 IRIS staff published systematic reviews of health effects of phthalates)[82,83,84,85]

[74] U.S. EPA. IRIS Glossary. https://www.epa.gov/iris/iris-glossary.
[75] Radke EG, Braun JM, Meeker JD, Cooper GS. Phthalate exposure and male reproductive outcomes: A systematic review of the human epidemiological evidence. Environ Int. 2018 Dec;121(Pt 1):764-793.
[76] Yost EE, Euling SY, Weaver JA, Beverly BEJ, Keshava N, Mudipalli A, Arzuaga X, Blessinger T, Dishaw L, Hotchkiss A, Makris SL. Hazards of diisobutyl phthalate (DIBP) exposure: A systematic review of animal toxicology studies. Environ Int. 2019 Apr;125:579-594.
[77] U.S. EPA (2023). Draft Proposed Approach for Cumulative Risk Assessment of High-Priority Phthalates and a Manufacturer-Requested Phthalate under the Toxic Substances Control Act, p. 102.
[78] U.S. EPA (2020). Risk evaluation for cyclic aliphatic bromide cluster (HBCD).
[79] U.S. EPA (2021). Human health toxicity values for perfluorobutane sulfonic acid and related compound potassium perfluorobutane sulfonate. https://cfpub.epa.gov/ncea/risk/recordisplay.cfm?deid=350888
[80] U.S. EPA (2020). Risk Evaluation for Percholorethylene (Ethene, 1,1,2,2-Tetrachloro-).
[81] U.S. EPA (2020). Risk Evaluation for Percholorethylene (Ethene, 1,1,2,2-Tetrachloro-).
[82] U.S. EPA (2020). Risk Evaluation for Trichloroethylene.
[83] U.S. EPA (2020). Risk Evaluation for Percholorethylene (Ethene, 1,1,2,2-Tetrachloro-).
[84] Radke EG, Braun JM, Meeker JD, Cooper GS. Phthalate exposure and male reproductive outcomes: A systematic review of the human epidemiological evidence. Environ Int. 2018 Dec;121(Pt 1):764-793.
[85] Yost EE, Euling SY, Weaver JA, Beverly BEJ, Keshava N, Mudipalli A, Arzuaga X, Blessinger T, Dishaw L, Hotchkiss A, Makris SL. Hazards of diisobutyl phthalate (DIBP) exposure: A systematic review of animal toxicology studies. Environ Int. 2019 Apr;125:579-594.

- acetylcholinesterase inhibition (numerous assessments of pesticides, including cumulative risk assessments of organophosphate and carbamate pesticides)[86,87]

EPA must either document that it has considered outcomes like altered thyroid hormone levels and other biochemical changes or cellular-level effects to be included in the animal and human evidence streams in the DIDP Draft Risk Evaluation, or provide a justification for why these outcomes should not be considered as potential hazards of DIDP.

Tagging biochemical and cellular-level outcomes as "supplemental, mechanistic," as directed in the PECO statement above, constrains the role of biochemical outcomes and other cellular changes to possibly providing biological support for apical outcomes, rather than considering precursors to apical outcomes as critical effects. Further, under EPA's proposed method, if no studies have been conducted of apical outcomes related to a biochemical outcome that has been studied, it is unclear whether the biochemical outcome will be considered at all. EPA says that supplemental studies "**may** be reviewed, evaluated for data quality, and incorporated into risk evaluations **as needed** for each chemical assessment"[88] (emphasis added), but it is unclear how a determination would be made to incorporate these studies into the risk evaluation, particularly in the absence of a related apical outcome study. Even if included to support a hazard conclusion based on apical outcomes, it appears that EPA rules out considering such studies for deriving a POD.

Exclusive reliance on studies of apical endpoints is also inconsistent with the best available science. An important theme of the NASEM 2007 *Toxicity Testing in the 21st Century* report was that toxicity testing should move away from reliance on testing of apical outcomes. Accordingly, EPA's research programs and other U.S. health agencies have invested heavily in this new direction. Government and academic toxicology labs now rarely conduct studies of apical endpoints because the science has shifted towards examining more sensitive endpoints representing upstream biological changes ("key events") that lead to apical outcomes. In addition, a restriction to consider only apical or organ-level studies may bias the evidence base of the TSCA risk evaluations toward inclusion of industry-funded guidelines studies, which are generally focused on apical endpoints.

d. EPA used multiple strategies to inappropriately exclude PECO-relevant health effects studies.

In past TSCA risk evaluations, EPA's practice was to exclude some health effects studies from consideration based on study quality evaluations; studies could be excluded based on a single perceived methodological shortcoming. EPA's draft systematic review protocol for DIDP says that, in response to recommendations from the NASEM, SACC and public comments, all relevant studies are included:

[86] U.S. EPA (2006). Organophosphorus cumulative risk assessment. https://www.regulations.gov/document/EPA-HQ-OPP-2006-0618-0002.
[87] U.S. EPA (2008). Revised N-methyl carbamate cumulative risk assessment. https://www.regulations.gov/document/EPA-HQ-OPP-2008-0347-0029.
[88] U.S. EPA (2021). Draft systematic review protocol supporting TSCA risk evaluations for chemical substances, p. 345.

> One main clarification is that *all references that undergo systematic review are considered for use in the risk evaluation*, even those that do not meet the various discipline and sub-discipline screening criteria or those that are categorized as supplemental information at title and abstract (TIAB) or full-text screening.[89]

This would be a welcome improvement to EPA's practice in TSCA risk evaluations; however, full consideration of EPA's systematic review procedures, as outlined in the draft protocol and hazard assessment, indicates that PECO-relevant health effects studies of DIDP can be excluded from the risk evaluation, and relevant studies were excluded through multiple procedures - some of which lack scientific justification.

First, the systematic review protocol for DIDP says EPA applied "further filtering" procedures to PECO-relevant health effects studies:

> References that met the PECO screening criteria and were categorized as having epidemiology information and/or animal toxicity information for the evaluation of human health hazard went through a fit-for-purpose further filtering step to determine which studies would move forward to data quality evaluation and data extraction.[90]

> To streamline the identification of studies containing potentially relevant data that had not previously been evaluated by an authoritative agency, modifications were implemented to the process described in the 2021 Draft Systematic Review Protocol…Following PECO-based screening, references that met PECO screening criteria for epidemiology underwent a two-step further filtering process to identify the subset of potentially relevant references that proceeded to data quality evaluation.[91]

> The main purpose of this further filtering step was to allow for the refinement of the references that would be considered for data quality evaluation and extraction.[92]

The protocol does not provide any explanation for why the application of the PECO was insufficient for determining studies to include in the DIDP Draft Risk Evaluation or why this "further filtering" process (which was not included in the 2021 TSCA draft systematic review method) was applied. It is also unclear why EPA found it necessary to "streamline" the process further when it was already extremely streamlined, with the most recent comprehensive literature search conducted in September 2019 and EPA's decision to expend very limited effort on pre-2018 epidemiology studies:

> Data quality evaluation and extraction wasn't conducted for any references published before 2018.[93]

[89] U.S. EPA (2024). Draft Risk Evaluation for Diisodecyl Phthalate (DIDP) – Systematic Review Protocol, p. 5.
[90] U.S. EPA (2024). Draft Risk Evaluation for Diisodecyl Phthalate (DIDP) – Systematic Review Protocol, p. 24.
[91] U.S. EPA (2024). Draft Risk Evaluation for Diisodecyl Phthalate (DIDP) – Systematic Review Protocol, p. 24.
[92] U.S. EPA (2024). Draft Risk Evaluation for Diisodecyl Phthalate (DIDP) – Systematic Review Protocol, p. 26.
[93] U.S. EPA (2024). Draft Risk Evaluation for Diisodecyl Phthalate (DIDP) – Systematic Review Protocol, p. 24.

Furthermore, it is unclear why a process to evaluate only 4 animal toxicology studies[94] needed additional filtering to exclude studies before assessing study quality.

Implementation of the further filtering step is also unclear. EPA provides a further filtering form for toxicology studies that includes a series of questions regarding the methods and outputs of a study. The form concludes with the Yes/No question "Should this reference move on to data extraction and evaluation?"[95] but no instructions are given for how the assessor is to answer this question.

EPA then explains that two of the four toxicology studies subjected to the further filtering procedure were excluded because "they were found to not provide quantitative data that could be used in our dose-response assessment."[96] Even if this is an accurate characterization of these studies, the statement indicates an inappropriate focus on quantitative information and the exclusion of relevant studies from informing hazard conclusions without scientific justification. The "further filtering" considerations are implicit amendments to the PECO statement that were not made available for public comment or peer review before the assessment was conducted, which is contrary to best practices for systematic review and contradicts EPA's claim that all relevant studies are considered in the DIDP Draft Risk Evaluation.

Second, studies that EPA deemed to be "uninformative" were not advanced to data extraction. The protocol states the EPA has continued its practice of excluding some studies based on study quality evaluations:

> Epidemiology and animal toxicity references with an overall quality determination (OQD) of High, Medium, or Low underwent data extraction; data wasn't extracted from Uninformative references.[97]

EPA's choice not to conduct data extraction for some studies based on the overall quality determination is equivalent to excluding these studies from the risk evaluation, again contradicting EPA's claim that all relevant studies are considered in the risk evaluation. Further, EPA's labeling of relevant studies as "Uninformative" is inappropriate and lacking in justification.

EPA never explains, in either the draft systematic review protocol for DIDP or the draft DIDP hazard assessment, how an OQD is derived from the study quality metrics. A statement at the end of the data quality evaluation forms for both epidemiology and toxicology studies indicates that EPA uses an automatic calculation of the OQD:

---

[94] U.S. EPA (2024). Draft Risk Evaluation for Diisodecyl Phthalate (DIDP) – Systematic Review Protocol, p. 26.
[95] U.S. EPA (2024). Draft Risk Evaluation for Diisodecyl Phthalate (DIDP) – Systematic Review Protocol, Table 4-1.
[96] U.S. EPA (2024). Draft Risk Evaluation for Diisodecyl Phthalate (DIDP) – Systematic Review Protocol, p. 27.
[97] U.S. EPA (2024). Draft Risk Evaluation for Diisodecyl Phthalate (DIDP) – Systematic Review Protocol, p. 60.

> Specify which OQD you would give this paper (either confirm the **auto calculated judgement** OR suggest a new one based on your professional judgement?[98] (emphasis added)

However, there is no other mention of "auto calculated judgement" in the protocol or hazard assessment. Further, there is no guidance given on when and with what basis an OQD not based on auto-calculation may be assigned.

In addition, EPA's continued use of the term "Uninformative" as an overall study rating is highly problematic. EPA's recent draft TSCA risk evaluation for formaldehyde demonstrates that an EPA determination of "Uninformative" is extremely unreliable and should not be used as a basis to exclude studies.[99]

For example, EPA's evaluation of study quality for oral toxicity studies of formaldehyde reveals the significant problems with assigning an OQD of "uninformative." EPA identified gastrointestinal effects as the most sensitive endpoint for oral exposure to formaldehyde. However, EPA classified the chronic oral exposure studies (by Til *et al*. and Tobe *et al*.) for gastrointestinal effects as "uninformative." After further consideration, EPA decided that these studies actually are informative, and that the Til *et al*. study should be used for dose-response analysis:

> Taken together, the three drinking water studies demonstrate a consistent pattern of gastrointestinal effects at comparable dose levels…While limitations in the two chronic drinking water studies resulted in OPPT data quality ratings of "uninformative for dose response" for the individual studies, the body of evidence across all three studies in combination increases the overall confidence in both the nature of the effects observed and the levels of formaldehyde exposure associated with those effects.[100]

> The three oral studies were selected to inform dose-response because they comprise the best available data on oral exposure to formaldehyde…when considered in conjunction with the other two studies, Til et al. 1989 contributes meaningful information to the WOE and dose-response despite the OPPT data quality rating of "uninformative."[101]

EPA's own analysis of its study quality ratings procedures therefore indicated that an overall study quality rating can be highly misleading and that labeling studies as "uninformative" or excluding studies based on the rating for a single study quality metric could erroneously lead to disregarding studies that constitute the best available science.

Accordingly, EPA must revise its approach to TSCA study quality evaluation to avoid disregarding studies based on pre-assigned labels that are unwarranted. By replacing the overall study quality determination with a domain- or metric-based approach, as the NASEM

[98] U.S. EPA (2024). Draft Risk Evaluation for Diisodecyl Phthalate (DIDP) – Systematic Review Protocol, Tables 5-5 and 5-7.
[99] U.S. EPA (2024). Draft Human Health Hazard Assessment for Formaldehyde.
[100] U.S. EPA (2024). Draft Human Health Hazard Assessment for Formaldehyde, pp. 30-31.
[101] U.S. EPA (2024). Draft Human Health Hazard Assessment for Formaldehyde, p. 32.

recommended for the TSCA program in 2021,[102] risk assessors can evaluate the ratings for each study in each domain at the evidence synthesis step to reach conclusions across the body of evidence, informed by the strengths and limitations of all relevant studies. These improved procedures should be applied to DIDP and are necessary for consistency with EPA's claim that all relevant studies are considered in the risk evaluation.

Finally, EPA appears to have also excluded studies from the DIDP Draft Risk Evaluation by other unexplained processes. Figure 4-6 of the draft systematic review protocol shows that out of 31 toxicology studies identified from previous hazard assessments of DIDP, 26 were excluded from consideration and only 5 studies were included in the draft risk evaluation.[103] No explanation is provided for the exclusion of these studies.

These examples demonstrate that EPA has not implemented procedures consistent with its claim that "*all references that undergo systematic review are considered for use in the risk evaluation."*[104] The TSCA systematic review process needs substantial revisions to correct a process that continues to exclude relevant evidence.

e. EPA continues to use unclear terminology regarding evidence synthesis and integration.

EPA's use of unclear terminology for evidence synthesis and integration is an additional scientific shortcoming of the approach to systematic review for DIDP. The NASEM has recommended the use of the term "evidence synthesis" for assembling the evidence and drawing conclusions from a single evidence stream (e.g. toxicology, epidemiology), and "evidence integration" for the subsequent process of drawing conclusions considering all evidence streams. The SACC review of EPA's 2021 Draft TSCA Method document reiterated this recommendation:

> The EPA did not follow the recommendation of NASEM to separate evidence synthesis from evidence integration. To quote NASEM: "Evidence synthesis deals with more homogeneous data within a single stream, and evidence integration deals with more heterogeneous data from multiple streams."[105]
>
> The EPA could improve the clarify, transparency, and efficiency of its process by adopting the NASEM recommendation to use "synthesis" for drawing conclusions separately for each evidence stream (i.e., human, animal, and mechanistic evidence) and use 'integration' for drawing conclusions considering all evidence streams in combination – in context of the risk evaluation process/needs.[106]

[102] NASEM (2021). The Use of Systematic Review in EPA's Toxic Substances Control Act Risk Evaluations, p. 36.
[103] U.S. EPA (2024). Draft Risk Evaluation for Diisodecyl Phthalate (DIDP) – Systematic Review Protocol, Figure 4-6, box 2a.
[104] U.S. EPA (2024). Draft Risk Evaluation for Diisodecyl Phthalate (DIDP) – Systematic Review Protocol, p. 5.
[105] U.S. EPA (2022). Science Advisory Committee on Chemicals Meeting Minutes and Final Report No. 2022-2, p. 83. https://www.regulations.gov/document/EPA-HQ-OPPT-2021-0414-0044.
[106] U.S. EPA (2022). Science Advisory Committee on Chemicals Meeting Minutes and Final Report No. 2022-2, p. 88. https://www.regulations.gov/document/EPA-HQ-OPPT-2021-0414-0044.

In the DIDP systematic review protocol, however, EPA disregards the advice of both the NASEM and the SACC by continuing to use the term "evidence integration" for both steps.[107] The draft DIDP hazard assessment further confuses matters by using the term "hazard identification"[108] instead of "evidence integration."

This is one more area in which EPA's approach differs from best practices in systematic review, violating the best available science requirement under TSCA.[109] In addition, failing to adopt consistent and vetted terminology decreases the clarity of the risk evaluation and creates confusion for peer reviewers and the public regarding the procedures applied to drawing conclusions from a single stream of evidence.

f. EPA's approach to evidence integration lacks clear procedures and clearly-stated conclusions regarding the hazards of DIDP.

EPA's DIDP systematic review protocol also fails to provide any clear indication of how integration of hazard evidence is conducted. For human health effects, the protocol section on evidence integration provides Table 6-2 regarding how to organize the evidence; however, this table does not address evidence integration and is instead concerned with the selection of studies for dose-response analysis.[110]

A further problem is that EPA does not apply consistent approaches to evidence integration. A key objective of the evidence integration process is to succinctly summarize the strength of the evidence concerning specific health endpoints and outcomes. This objective is advanced by pre-specifying a standard set of evidence descriptors. For example, EPA's IRIS program uses the terms "evidence demonstrates," "evidence indicates," and "evidence suggests" as hazard conclusions. No such terms are used in TSCA risk evaluations, except in instances like the formaldehyde draft risk evaluation, where conclusions from IRIS are reported. In the DIDP Draft Risk Evaluation, EPA uses a range of ambiguous phrases in its hazard conclusions, including:

- "strong evidence" (developmental effects and liver toxicity)[111]
- "consistent evidence" (kidney toxicity)[112]
- "some limited evidence" (neurotoxicity)[113]
- "some evidence" (immune system toxicity).[114]

Without further standardization and definition of terms, it is difficult for readers to gain a clear, concise understanding of EPA's hazard conclusions. It is unclear, for example, if "consistent

[107] U.S. EPA (2024). Draft Risk Evaluation for Diisodecyl Phthalate (DIDP) – Systematic Review Protocol, pp. 101-104.
[108] U.S. EPA (2024). Draft Human Health Hazard Assessment for Diisodecyl Phthalate (DIDP), p. 23.
[109] 15 U.S.C. § 2625(h).
[110] U.S. EPA (2024). Draft Human Health Hazard Assessment for Diisodecyl Phthalate (DIDP) p.54, Table 6-2.
[111] U.S. EPA (2024). Draft Human Health Hazard Assessment for Diisodecyl Phthalate (DIDP), p. 27 and p. 31.
[112] U.S. EPA (2024). Draft Human Health Hazard Assessment for Diisodecyl Phthalate (DIDP), p. 35.
[113] U.S. EPA (2024). Draft Human Health Hazard Assessment for Diisodecyl Phthalate (DIDP), p. 39.
[114] U.S. EPA (2024). Draft Human Health Hazard Assessment for Diisodecyl Phthalate (DIDP), p. 42.

evidence" is equivalent to "strong evidence," or whether "some limited evidence" is equivalent to "some evidence."

EPA should apply a consistent procedure for evidence integration for all endpoints, including a pre-specified set of descriptors that are considered for each endpoint.

g. EPA released an incomplete draft systematic review protocol for DIDP that was not released in advance of the draft risk evaluation.

Along with the DIDP Draft Risk Evaluation, EPA released a chemical-specific systematic review protocol as a supplemental file. Publication of a chemical-specific protocol is consistent with best available scientific methods in systematic review and responds to recommendation of the NASEM and the SACC. However, the comments above demonstrate many flaws and deficiencies in the protocol and the procedures applied to conducting the risk evaluation. Public release of the protocol for public comment and peer review in advance of conducting the risk evaluation would have provided an opportunity for early identification and correction of the many critical deficiencies described above. For future TSCA risk evaluations, EPA must publish a chemical-specific systematic review protocol for public comment before completing the draft risk evaluation, as recommended by the Institute of Medicine and the NASEM as a best practice for systematic review.[115,116]

The TSCA program should follow the established procedures of EPA's IRIS program, which makes a draft protocol for each assessment publicly available in advance of its release for public comment. Following the public comment process, the IRIS program then publishes an updated protocol, as needed. For example, for the IRIS assessments of five per- and polyfluoroalkyl substances ("PFAS"), a draft protocol was made available for public comment for 45 days. The IRIS program then followed up with a revised protocol to address public comments, with documentation of the changes, that was published before the release of the PFAS draft assessments.[117] EPA should be following this same approach for all TSCA risk evaluations.

h. EPA should prepare a new TSCA systematic review handbook that is aligned with the best available scientific methods and issue updated draft systematic review protocols for all risk evaluations currently in development.

To adhere to best practices in systematic review, including those recommended by the NASEM and SACC, EPA should issue a new TSCA systematic review methodology document that states methods to be applied consistently to all TSCA risk evaluations. EPA should also prepare a chemical-specific systematic review protocol for each TSCA risk evaluation it conducts, and these protocols should be complete, stand-alone documents that do not refer to the 2021 Draft TSCA Method for critical elements. The chemical-specific protocols for ongoing and future risk evaluations should also be released for public comment well before the draft risk evaluations are completed to allow for public input, scrutiny, and opportunities for improvement. We urge EPA

[115] Institute of Medicine (2011). Finding what works in health care: Standards for systematic reviews.
[116] National Research Council (2014). Review of EPA's Integrated Risk Information System (IRIS) process.
[117] U.S. EPA (2021). Systematic Review Protocol for the PFAS IRIS Assessments. https://cfpub.epa.gov/ncea/iris_drafts/recordisplay.cfm?deid=345065 (accessed 1 February 2024).

to consistently adopt the practices of the IRIS program for systematic review protocol development and publication across all EPA programs and offices.

4. EPA's occupational and consumer exposure assessments for DIDP are not consistent with the best available science.

a. EPA's determination of unreasonable risk in occupational settings discounts and disregards EPA's own occupational risk estimates for non-cancer effects.

In previous TSCA risk evaluations, EPA has typically determined whether a condition of use for a particular chemical contributes to unreasonable risk through comparison to benchmark values. For non-cancer effects, the comparison is to a benchmark MOE that is based on selection of applicable uncertainty factors. If the MOE for a particular exposure scenario, calculated as the POD divided by the estimated human exposure, is less than the identified benchmark MOE, EPA has typically concluded that the exposure constitutes an unreasonable risk. For example, the conditions of use identified by EPA as the supporting basis for the final TSCA unreasonable risk determination for TCE based on non-cancer effects to workers and consumers correspond exactly to the exposure scenarios in which the calculated MOEs are lower than the benchmark MOEs[118].

In the DIDP Draft Risk Evaluation, EPA says:

> A calculated MOE that is less than the benchmark MOE is a starting point for informing a determination of unreasonable risk of injury to health, based on non-cancer effects[119].
>
> When determining whether a chemical substance presents unreasonable risk to human health or the environment, **calculated risk estimates are not "bright-line" indicators of unreasonable risk**, and **EPA has the discretion to consider other risk-related factors** in addition to risks identified in the risk characterization. (emphasis added)[120]

This interpretation of the MOE provided for DIDP is significantly different from what was stated in previous TSCA risk evaluations. EPA's 2023 draft supplement to the risk evaluation for 1,4-dioxane stated that "[t]he MOE estimate is interpreted as **indicating a human health risk** if the MOE estimate is less than the benchmark MOE;"[121] similarly, the 2020 final risk evaluation for methylene chloride says "The MOE estimate was **interpreted as a human health risk** if the MOE estimate was less than the benchmark MOE"[122] (emphasis added).

In addition, EPA's claim that it does not use the MOE as a bright line indicator is false, as the draft risk evaluation states that:

---

[118] U.S. EPA (2022). Risk Evaluation for Trichloroethylene. Final Revised Unreasonable Risk Determination for Trichloroethylene, Tables 5-1 and 5-2. https://www.epa.gov/system/files/documents/2023-01/TCE_Final%20Revised%20RD_12-21-22-FINAL-v2.pdf.
[119] U.S. EPA (2024). Draft Risk Evaluation for Diisodecyl Phthalate (DIDP), p. 161.
[120] U.S. EPA (2024). Draft Risk Evaluation for Diisodecyl Phthalate (DIDP), p. 113.
[121] U.S. EPA (2023). Draft Supplement to the Risk Evaluation for 1,4-Dioxane. p 136.
[122] U.S. EPA (2020). Risk Evaluation for Methylene Chloride (Dichloromethane, DCM), p. 365.

> if the MOE estimate is equal to or exceeds the benchmark MOE, the risk is not considered to be of concern and mitigation is not needed.[123]

In multiple cases for DIDP, EPA's own MOE calculations indicate unreasonable risk, but EPA dismisses these calculations, claiming that they are overestimates without presenting any supporting evidence. In the DIDP Draft Risk Evaluation, EPA found that only one COU presents an unreasonable risk to human health despite 9 other COUs having MOEs less than the benchmark MOE. In EPA's previous risk evaluations, these results would have translated into a determination that all 10 occupational COUs contribute to unreasonable risks from DIDP based on non-cancer risks. Table 1 highlights the numerous instances where EPA disregarded or dismissed, without scientific justification, risks greater than its already under-protective and flawed benchmarks.

Table 1- EPA's scientifically unsupported rationale for dismissing risks for COUs where the calculated MOE was less than the benchmark MOE.

COU	OES	Calculated MOE (Benchmark MOE=30)	EPA's explanation for dismissing risk
Plastics material and resin manufacturing	PVC plastics compounding	aggregate high-end worker (average adult) acute exposure was **25** aggregate high-end worker (female of reproductive age) acute exposure was **24**	"exposure and risk estimates are based on the assumption that the concentration of DIDP in workplace dust is the same as the concentration of DIDP in PVC plastics or non-PVC materials, respectively. However, it is likely that workplace dust contains a variety of constituents and that the concentration of DIDP in workplace dust is less than the concentration of DIDP in PVC or non-PVC products. Therefore, central tendency values of exposure are expected to be more reflective of true worker exposures"[124]
Other (part of the formulation for manufacturing synthetic leather)	PVC plastics compounding	aggregate high-end worker (average adult) acute exposure was **25** aggregate high-end worker (female of reproductive age) acute exposure was **24**	"exposure and risk estimates are based on the assumption that the concentration of DIDP in workplace dust is the same as the concentration of DIDP in PVC plastics or non-PVC materials, respectively. However, it is likely that workplace dust contains a variety of constituents and that the concentration of DIDP in workplace dust is less than the concentration of DIDP in PVC or non-PVC products. Therefore, central tendency values of exposure are expected to be more reflective of true worker exposures"[125]

[123] U.S. EPA (2024). Draft Risk Evaluation for Diisodecyl Phthalate (DIDP), p. 113.
[124] U.S. EPA (2024). Draft Risk Evaluation for Diisodecyl Phthalate (DIDP), p. 117.
[125] U.S. EPA (2024). Draft Risk Evaluation for Diisodecyl Phthalate (DIDP), p. 117.

Abrasives manufacturing	Application of adhesives and sealants	high-end worker (average adult) acute, intermediate, and chronic inhalation exposure ranged from **3.3 to 4.8** high-end worker (female of reproductive age) acute, intermediate, and chronic inhalation exposure ranged from **2.9 to 4.3** aggregate high-end worker (average adult) acute, intermediate, and chronic exposure ranged from **3.2 to 4.6** aggregate high-end worker (female of reproductive age) acute, intermediate, and chronic exposure ranged from **2.9 to 4.2**	"high-end inhalation exposure estimates are more representative of high-pressure spray applications whereas the central tendency estimates are more representative of low-pressure applications"[126]
Adhesives and sealants (including plasticizers in adhesives and sealants)	Application of adhesives and sealants	high-end worker (average adult) acute, intermediate, and chronic inhalation exposure ranged from **3.3 to 4.8** high-end worker (female of reproductive age) acute, intermediate, and chronic inhalation exposure ranged from 2**.9 to 4.3** aggregate high-end worker (average adult) acute, intermediate, and chronic exposure ranged from **3.2 to 4.6** aggregate high-end worker (female of reproductive age) acute, intermediate, and chronic exposure ranged from **2.9 to 4.2**	"high-end inhalation exposure estimates are more representative of high-pressure spray applications whereas the central tendency estimates are more representative of low-pressure applications"[127]
Lacquers, stains, varnishes, and floor finishes (as plasticizers)	Application of adhesives and sealants	high-end worker (average adult) acute, intermediate, and chronic inhalation exposure ranged from **3.3 to 4.8** high-end worker (female of reproductive age) acute, intermediate, and chronic inhalation exposure ranged from **2.9 to 4.3**	"high-end inhalation exposure estimates are more representative of high-pressure spray applications whereas the central tendency estimates are more representative of low-pressure applications"[128]

[126] U.S. EPA (2024). Draft Risk Evaluation for Diisodecyl Phthalate (DIDP), p. 114.
[127] U.S. EPA (2024). Draft Risk Evaluation for Diisodecyl Phthalate (DIDP), p. 114.
[128] U.S. EPA (2024). Draft Risk Evaluation for Diisodecyl Phthalate (DIDP), p. 114.

		aggregate high-end worker (average adult) acute, intermediate, and chronic exposure ranged from **3.2 to 4.6** aggregate high-end worker (female of reproductive age) acute, intermediate, and chronic exposure ranged from **2.9 to 4.2**	
Paints and coatings (including surfactants in paints and coatings)	Application of paints and coatings	aggregate high-end worker (average adult) acute exposure was **24** aggregate high-end worker (female of reproductive age) acute exposure was **23**	"high-end inhalation exposure estimates are more representative of high-pressure spray applications whereas the central tendency estimates are more representative of low-pressure applications"[129]
Lacquers, stains, varnishes, and floor finishes (as plasticizers)	Application of paints and coatings	aggregate high-end worker (average adult) acute exposure was **24** aggregate high-end worker (female of reproductive age) acute exposure was **23**	"high-end inhalation exposure estimates are more representative of high-pressure spray applications whereas the central tendency estimates are more representative of low-pressure applications"[130]
Ink, toner, and colorant products	Application of paints and coatings	aggregate high-end worker (average adult) acute exposure was **24** aggregate high-end worker (female of reproductive age) acute exposure was **23**	"high-end inhalation exposure estimates are more representative of high-pressure spray applications whereas the central tendency estimates are more representative of low-pressure applications"[131]
Inspection fluid/penetrations	Use of penetrants and inspection fluids	high-end worker (average adult) acute, intermediate, and chronic inhalation exposure ranged from **13 to 19** high-end worker (female of reproductive age) acute, intermediate, and chronic inhalation exposure ranged from **12 to 17** aggregate high-end worker (average adult) acute, intermediate, and chronic exposure ranged from **11 to 17** aggregate high-end worker (female of reproductive age) acute, intermediate, and chronic exposure ranged from **11 to 16**	"Aerosol application may overestimate inhalation exposures for brush application methods. Therefore, the central tendency exposure levels are expected to be representative of the commercial **COU: 'Inspection fluid/penetrant'** due to uncertainties in both product concentration and method of application."[132]

[129] U.S. EPA (2024). Draft Risk Evaluation for Diisodecyl Phthalate (DIDP), p. 114.
[130] U.S. EPA (2024). Draft Risk Evaluation for Diisodecyl Phthalate (DIDP), p. 114.
[131] U.S. EPA (2024). Draft Risk Evaluation for Diisodecyl Phthalate (DIDP), p. 114.
[132] U.S. EPA (2024). Draft Risk Evaluation for Diisodecyl Phthalate (DIDP), p. 116.

It is also concerning that EPA chose to disregard high-end risk estimates at the final stages of risk determination, only after finding that risks are high for the 9 COUs highlighted in Table 1. While EPA mentions data uncertainties in the Draft Environmental Release and Occupational Exposure Assessment Document, EPA also concluded that the weight of the scientific evidence for all occupational exposures estimates was "moderate" and the estimates were "plausible." EPA does not provide sufficient evidence in the DIDP Draft Risk Evaluation for its claims that the high-end estimates are not representative of exposures for at least some workers. EPA further does not present evidence justifying the use of central tendency estimates to characterize exposures and risks to all workers in each COU. Ignoring calculated risk at the final stage of the draft risk evaluation based on flawed and scientifically unsupported rationale undermines the integrity of the risk estimates, and suggests that EPA is disregarding high-end estimates solely to avoid determining a contribution to unreasonable risk for each occupational COU. EPA must adopt a more transparent, consistent, and accountable approach to risk assessment. In any instance where EPA characterizes risk for a COU based only on central tendency estimates, EPA must provide sufficient evidence demonstrating that the central tendency does not underestimate exposure and risk for more-exposed individuals or susceptible subpopulations. Uncertainties identified by EPA must be addressed early in the exposure assessment, and all reasonably foreseeable exposures must be considered and accounted for when evaluating COUs.

In addition, EPA's attempts to justify disregarding the high-end estimates, including repeated mentions of uncertainties and lack of data, indicate that EPA failed in its obligation to ensure that it obtained the necessary data needed to conduct a defensible risk evaluation. This is particularly concerning for a manufacturer-requested risk evaluation, where, according to the preamble to the original risk evaluation framework rule (which was in place at the time EPA granted the request), the

> manufacturers are required to submit all the information necessary to complete risk evaluation[133]

Further, according to the framework rule, EPA should have initiated the risk evaluation only if it had obtained the necessary data from the manufacturers:

> EPA will grant the request if it determines that… EPA has the required information necessary for conducting a risk evaluation… Bases for a denial, include the manufacturer has not provided sufficient information to complete the risk evaluation[134].

Even having granted the manufacturer request without adequate data, there are no indications that EPA utilized its authority under TSCA to obtain data after initiating the DIDP Draft Risk Evaluation. Given the potentially significant data gaps, EPA's high-end exposure estimates make appropriate use of the reasonably available data and should be used as a basis for unreasonable risk determination.

[133] 82 FR 33726.
[134] 40 CFR § 702.37(e)(6).

b. EPA inappropriately disregards high-end exposure estimates without justification for its unreasonable risk determinations for workers, ignoring variability in exposures and TSCA's requirement to assess risks to groups with greater exposures.

The practice of utilizing high-end exposure estimates is scientifically well-supported and is consistent with both the requirements of TSCA and previous TSCA risk evaluations. This approach is crucial for ensuring that the risk evaluation comprehensively addresses all potential risks, particularly to the most vulnerable and highly exposed groups within the workforce.

However, in the DIDP Draft Risk Evaluation, EPA applies the central tendency estimates instead of high-end exposure estimates, effectively discounting any unreasonable risks to workers. This raises significant concerns about the adequacy and methods of the risk evaluation. The justification provided by the EPA for preferring central tendency estimates—stating that high-end exposure estimates are less representative of occupational conditions of use (COUs)—lacks sufficient evidence. EPA cites that high-end inhalation exposure estimates typically represent high-pressure spray applications and suggests that central tendency estimates are more reflective of low-pressure applications, including non-spray methods.[135] However, EPA presents no evidence to support the notion that high-end exposures are an overestimation or that such exposure scenarios are unlikely to occur. Moreover, EPA's sole reliance on central tendency estimates likely underestimates exposures in scenarios that do not conform to this median.

More critically, the use of central tendency estimates fails to consider the risk to individuals exposed at levels above this median, potentially disregarding the health risks to half of the exposed population. This approach does not align with TSCA's mandate to identify and protect potentially exposed or susceptible subgroups (PESS), characterized by greater exposure levels than the general population.[136]

Applying only central tendency estimates for the risk evaluation also means that EPA will potentially overlook significant risks, particularly for workers engaged in high-exposure tasks or those exposed to multiple chemical and non-chemical stressors. Special consideration should be given to more vulnerable workers, including women of reproductive age and other PESS, who might face heightened risks even at lower levels of exposure.

To adhere to the requirements of TSCA and to ensure a robust protection for all workers, the EPA should employ high-end exposure estimates that represent at least the 95th percentile of exposure—preferably even higher, such as the 99th percentile. This adjustment is necessary to accurately reflect the risk for the most exposed individuals and to ensure that all COUs are evaluated with an appropriate level of concern, particularly those currently deemed as less certain or not contributing to unreasonable risk.[137]

[135] U.S. EPA (2024). Draft Risk Evaluation for Diisodecyl Phthalate (DIDP), p. 114.
[136] 15 U.S.C. §2602(12).
[137] 15 U.S.C. §2602(12).

c. EPA's consumer exposure assessment disregards unreasonable risk to more vulnerable groups, such as infants and toddlers, with little to no scientific justification.

EPA's risk characterization for consumer exposures found one consumer COU that had an acute non-cancer MOE less than the benchmark MOE of 30:

> COU: **Consumer Uses: Packaging, paper, plastic, hobby products: Plastic and rubber products (textiles, apparel, and leather; vinyl tape; flexible tubes; profiles; hoses),** the aggregate acute high-end exposure estimates for wallpaper are 27 and 28 for infants and toddlers, respectively.[138]

However, EPA chose to disregard high-end risk estimates in the unreasonable risk determination for this COU after finding risks of concern. Neither the DIDP Draft Consumer and Indoor Dust Exposure Assessment nor the DIDP Draft Risk Evaluation provide sufficient evidence for why this high-end exposure was dismissed. EPA dismisses the risk presented to infants and toddlers for this COU citing that:

> **The high-intensity model conservatively assumes that a relatively large surface area of the house is covered with in-place wallpaper (200 m2), a DIDP weight fraction of 0.26 percent** (based on two wallpaper samples containing both DINP and DIDP that was reported in 2001 study of four PVC wallpapers), and **the infant stays at home all day long**.(emphasis added)[139]

EPA's rationale is not scientifically supported and concludes that the model assumptions are conservative without any supporting basis. EPA provides no reason to doubt that a home might have 200 m^2 of wallpaper, especially a new home or a home that was recently remodeled in anticipation of the birth of an infant. A small number of samples does not indicate any bias in the weight fraction – the value obtained from two samples could be either an overestimate or an underestimate, and there is no basis to conclude that this value is conservative. Further, it is not unusual for infants to stay home all day long.[140]

By only relying on central tendency exposure estimates, EPA has likely underestimated consumer risks by not accounting for scenarios in which higher indoor air concentrations may occur. Again, this approach raises concerns about the thoroughness and credibility of the risk assessment conducted. Ignoring calculated risks at the final stage of the draft risk evaluation undermines the integrity of the risk estimates, suggesting that EPA is using the central tendency to avoid making an unreasonable risk determination for consumer COUs.

5. EPA failed to adequately identify potentially exposed or susceptible subpopulations (PESS), as required by TSCA.

[138] U.S. EPA (2024). Draft Risk Evaluation for Diisodecyl Phthalate (DIDP), p. 139.
[139] U.S. EPA (2024). Draft Risk Evaluation for Diisodecyl Phthalate (DIDP), p. 167.
[140] U.S. EPA. Child-Specific Exposure Factors Handbook (2008, Final Report). U.S. Environmental Protection Agency, Washington, DC, EPA/600/R-06/096F, 2008.

EPA has failed to meet its requirement under TSCA to identify, consider, and account for risk to "potentially exposed or susceptible subpopulations" ("PESS") in the DIDP Draft Risk Evaluation.[141] EPA excluded multiple potential PESS and among the PESS identified, EPA did not apply a transparent methodology for quantifying the risk of harm to each identified PESS using the best available science. This omission is consistent with previous risk evaluations where EPA regularly underestimated the risk to PESS due to a lack of adequate identification and consideration of PESS. By not adequately considering PESS, EPA is violating TSCA's requirements. EPA therefore must adopt a consistent framework for identifying and quantifying the risk of harm to PESS from DIDP exposures.

Identification and consideration of PESS for each chemical assessed is a critical aspect of conducting risk evaluation under TSCA, as TSCA requires EPA to

> determine whether a chemical substance presents an unreasonable risk of injury to health or the environment, without consideration of costs or other nonrisk factors, including an unreasonable risk to a potentially exposed or susceptible subpopulation.[142]

In the final 2024 TSCA Risk Evaluation Framework Rule, EPA defined PESS as:

> Potentially exposed or susceptible subpopulation means a group of individuals within the general population identified by EPA who, due to either greater susceptibility or greater exposure, may be at greater risk than the general population of adverse health effects from exposure to a chemical substance or mixture, such as infants, children, pregnant women, workers, the elderly, or overburdened communities.[143]

EPA needs to develop and apply a consistent approach to identify all PESS. To date, EPA has not employed a consistent or structured approach to identifying PESS in its TSCA risk evaluations, including scope documents for ongoing risk evaluations. EPA's approach and terminology for identifying PESS varied considerably in the first 10 risk evaluations. These inconsistencies include: differences in whether health conditions related to a chemical's hazards were considered in identifying PESS; and whether fenceline communities were included as PESS.[144] To remedy the problem of inconsistent and incomplete identification of PESS, Rayasam *et al.* recommended that:

> EPA should prepare a comprehensive methodology to identify PESS and quantify their risks consistently within and across the TSCA risk evaluations.[145]

---

[141] 15 U.S.C. §2605(b)(4)(A).
[142] 15 U.S.C. §2605(b)(4)(A).
[143] U.S. EPA (2024). Procedures for Chemical Risk Evaluation Under the Toxic Substances Control Act, § 702.33.
[144] Rayasam, S. D. G., Koman, P. D., Axelrad, D. A., Woodruff, T. J., Chartres, N. (2022). Toxic Substances Control Act (TSCA) Implementation: How the Amended Law Has Failed to Protect Vulnerable Populations from Toxic Chemicals in the United States. Environmental science & technology, 56(17), 11969–11982. https://doi.org/10.1021/acs.est.2c02079.
[145] Rayasam, S. D. G., Koman, P. D., Axelrad, D. A., Woodruff, T. J., Chartres, N. (2022). Toxic Substances Control Act (TSCA) Implementation: How the Amended Law Has Failed to Protect Vulnerable Populations from Toxic Chemicals in the United States. Environmental science & technology, 56(17), 11969–11982. https://doi.org/10.1021/acs.est.2c02079].

EPA has not yet proposed such a methodology. While the listing of potential PESS in Table 7-1 in the DIDP Draft Risk Evaluation is a useful initial step towards developing a consistent, structured approach to identifying PESS in TSCA risk evaluations,[146] EPA has taken a step backwards with the exclusion of more detailed evaluations of PESS based on both greater exposure and greater susceptibility that were included in recent risk evaluations. Table 7-1 gives explicit consideration to each of the following: lifestage, pre-existing disease or disorder, lifestyle activities, socio-demographic factors, nutrition, genetics/epigenetics, and other chemical and non-chemical stressors, yet EPA fails to fully consider all PESS within each category identified.[147] EPA also failed to identify PESS categories that were identified in previous risk evaluations, including geographic factors, effectively taking another step backwards with PESS identification.

In addition, EPA's evaluation and application of uncertainty factors aimed at protecting PESS falls short at every step and are insufficient for protecting PESS. EPA quantitatively adjusted for differences in human susceptibility only with the application of the standard human variability uncertainty factor of 10X. However, the WHO and other authoritative bodies have demonstrated that even the traditional 10X uncertainty factor is insufficient for fully accounting for risk in sensitive groups and recommend the use larger uncertainty factors.[148] Instead of increasing the use of uncertainty factors to account for the wide range of vulnerability and variability in the human population, EPA uses inadequate default uncertainty factors, which will result in an underestimation of risk, particularly for PESS.

For the identified PESS, EPA also concluded that, due to a lack of chemical specific data for each PESS, no further adjustment is necessary. TSCA does not require chemical-specific quantitative data to identify or evaluate risks to PESS. Instead, TSCA requires EPA to rely on the "best available science" when evaluating risks to PESS. The best available science demonstrates that both intrinsic factors, which include biological traits like age, genetic makeup, and pre-existing health conditions, and extrinsic factors, which include psychosocial stress from experiencing income inequality, violence, racism, healthcare inequity, or food insecurity, can individually or collectively increase susceptibility to harm from chemical exposures.[149]

---

[146] U.S. EPA (2024). Draft Human Health Hazard Assessment for Diisodecyl Phthalate (DIDP), p. 61, Table 7-1.
[147] U.S. EPA (2024). Draft Human Health Hazard Assessment for Diisodecyl Phthalate (DIDP), p. 61, Table 7-1.
[148] Varshavsky, J. R., Rayasam, S. D. G., Sass, J. B., Axelrad, D. A., Cranor, C. F., Hattis, D., Hauser, R., Koman, P. D., Marquez, E. C., Morello-Frosch, R., Oksas, C., Patton, S., Robinson, J. F., Sathyanarayana, S., Shepard, P. M., & Woodruff, T. J. (2023). Current practice and recommendations for advancing how human variability and susceptibility are considered in chemical risk assessment. Environmental Health, 21(Suppl 1), 133. https://doi.org/10.1186/s12940-022-00940-1.
[149] Woodruff, T. J., Rayasam, S. D. G., Axelrad, D. A., Koman, P. D., Chartres, N., Bennett, D. H., Birnbaum, L. S., Brown, P., Carignan, C. C., Cooper, C., Cranor, C. F., Diamond, M. L., Franjevic, S., Gartner, E. C., Hattis, D., Hauser, R., Heiger-Bernays, W., Joglekar, R., Lam, J., … Zeise, L. (2023). A science-based agenda for health-protective chemical assessments and decisions: Overview and consensus statement. Environmental Health,21(1), 132. https://doi.org/10.1186/s12940-022-00930-3; Rachel Morello-Frosch et al., Understanding the Cumulative Impacts of Inequalities in Environmental Health: Implications for Policy, 30 Health Affs. 879 (2011), https://www.healthaffairs.org/doi/pdf/10.1377/hlthaff.2011.0153; Cliona M. McHale et al., Assessing Health Risks

EPA should therefore focus first on identifying susceptible subpopulations based on either chemical-specific evidence or the broader literature on intrinsic and extrinsic susceptibility factors, and then, as a separate step, consider how to adequately account for the elevated risks for each group, in some cases by using scientifically-supported uncertainty factors. The initial identification of PESS, however, should not be contingent on chemical-specific data. Once the appropriate groups are identified as PESS, EPA should then consider the availability of chemical-specific data. When such data are absent, the application of appropriate adjustment factors (beyond the customary 10x factor for human variability) should be applied to ensure that risks to PESS are not underestimated.[150] Table 2 describes the PESS considerations listed in the DIDP Draft Risk Evaluation, the gaps in PESS identification or consideration, and recommended science-based uncertainty factors that should be employed to fully account for risk posed to each group.

Table 2. PESS considerations and recommended uncertainty factors.

PESS category	PESS identified by EPA	PESS EPA failed to consider	EPA Proposed UF	PRHE Recommended UF
Lifestage	Embryos/fetuses, Infants and children, females of reproductive age/pregnancy/lactating status, males of reproductive age, children, elderly	Did not sufficiently account for variability and vulnerability of certain life stages in the human population	No additional UFs beyond the 10X identified for general human variability: POD for developmental endpoints is thought to be protective of effects in offspring, children, and females of reproductive	42X and additional 10X for susceptible life stages, including pregnant women

from Multiple Environmental Stressors: Moving from G×E to I×E, 775 Mutational Rsch. 11 (2018), https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5863617/; Devon C. Payne-Sturges et al., Methods for Evaluating the Combined Effects of Chemical and Nonchemical Exposures for Cumulative Environmental Health Risk Assessment, 15 Int'l. J. Env't Rsch. & Pub. Health 2797 (2018), https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6313653/; Gilbert C. Gee et al., Environmental Health Disparities: A Framework Integrating Psychosocial and Environmental Concepts, 112 Env't Health Persps. 1645 (2004), https://doi.org/10.1289/ehp.7074; Gina M. Solomon et al., Cumulative Environmental Impacts: Science and Policy to Protect Communities 37 Ann. Rev. Pub. Health 83, 87–88 (2016), https://www.annualreviews.org/doi/pdf/10.1146/annurev-publhealth-032315-021807; Patricia D. Koman et al., Population Susceptibility: A Vital Consideration in Chemical Risk Evaluation Under the Lautenberg Toxic Substances Control Act, 17 PLoS Biology 1, 4 (2019), https://journals.plos.org/plosbiology/article?id=10.1371/

[150] Varshavsky et al. Current Practice and Recommendations for Advancing How Human Variability and Susceptibility Are Considered in Chemical Risk Assessment, 21(Suppl 1) Env't Health Article No. 133, at 3 (2023), https://doi.org/10.1186/s12940-022-00940-1.

			age/pregnancy/ lactating status	
Pre-existing disease or disorder	Health outcomes/target organs, toxicokinetics	Did not sufficiently account for variability in the human population	No additional UFs beyond the 10X identified for general human variability	42X and an additional 10X for pre-existing disease
Lifestyle activities	smoking, alcohol consumption, physical activity	Failed to adjust for lifestyle activities	Only a "qualitative discussion"	42X and an additional 10X for non-chemical stressors
Socio-demographic factors	Race/ethnicity Socio-economic status, sex/gender	Failed to adjust for SES factors and race/ethnicity	Only a "qualitative discussion" for race/ethnicity No adjustment for SES No additional UFs beyond the 10X identified for general human variability for sex/gender	42X and an additional 10X for non-chemical stressors
Nutrition	Diet, malnutrition	Failed to adjust for nutrition factors	Only a "qualitative discussion"	42X and an additional 10X for non-chemical stressors
Genetics/epigenetics	Target organs, toxicokinetics.	Did not sufficiently account for the genetic variability in human populations	Only a "qualitative discussion" for target organs No additional UFs beyond the 10X identified for general human	42X

			variability for toxicokinetics	
Other chemical and non-chemical stressors	Built environment, social environment, chemical co-exposures	Failed to adjust for other chemical and non-chemical stressors	Only a "qualitative discussion"	42X and an additional 10X for multiple chemical and non-chemical stressors
Geographic factors	Not Applicable	EPA failed to identify geographic factors as a PESS category	Not Applicable	42X and an additional 10X for non-chemical stressors
Unique activities	Not Applicable	EPA failed to identify unique activities as a PESS category	Not Applicable	42X and an additional 10X for non-chemical stressors

6. EPA failed to conduct a background exposure assessment, underestimating risk to potentially exposed or susceptible subpopulations.

Phthalates such as DIDP have become ubiquitous contaminants worldwide to which the general population is commonly exposed through multiple pathways, including water, air, and inhalation and/or ingestion of household dust.[151] DIDP is primarily used as a plasticizer to make flexible polyvinyl chloride (PVC). It is also used to make building and construction materials, automotive care and fuel products, and other commercial and consumer products such as adhesives, sealants, paints, coatings, and electrical products.[152] DIDP is also found in several common household items such as food packaging materials, nail polishes, fragrances, and pharmaceuticals.[153]

However, EPA failed to account for these multiple sources of exposure in the DIDP Draft Risk Evaluation. Instead, EPA stated that certain significant pathways of exposure to the general population, including food, food packaging materials, nail polishes, and fragrances, could not be considered because they constitute "non-TSCA" uses.[154] EPA's rationale for this decision is that these other pathways of exposure will be assessed and managed by statutes such as the Clean Air Act and the Federal Food, Drug, and Cosmetic Act. However, exposures via these pathways are highly relevant and reasonably foreseeable across the human population, and cannot be excluded when evaluating the human health risks posed by DIDP. EPA is required under TSCA to account

[151] U.S. EPA (2024). Draft Risk Evaluation for Diisodecyl Phthalate (DIDP), p. 10.
[152] U.S. EPA (2024). Draft Risk Evaluation for Diisodecyl Phthalate (DIDP), p. 10.
[153] U.S. EPA (2024). Draft Risk Evaluation for Diisodecyl Phthalate (DIDP), p. 10.
[154] U.S. EPA (2024). Draft Risk Evaluation for Diisodecyl Phthalate (DIDP), p. 10.

for all "reasonably foreseeable" pathways of exposure.[155] EPA must also conduct risk evaluations using "scientific information, technical procedures, measures, methods, protocols, methodologies, or models, employed in a manner consistent with the best available science."[156] The NASEM recommends consideration of background exposures when conducting a risk evaluation for both individual chemicals and categories of chemicals through a cumulative risk assessment,[157] citing that background exposures at "even small doses may have a relevant biological effect."[158]

Given the widespread exposure to DIDP across the general population and susceptible populations through food, plastic food storage products, nail polishes, and other "non-TSCA" uses, the failure to consider exposures from those uses would be contrary to TSCA's requirements to consider all reasonably foreseeable exposure pathways and to identify and address risks to PESS. While EPA may not be able to directly regulate some uses under TSCA, EPA cannot adequately evaluate the conditions of use that are subject to TSCA regulation or control their unreasonable risks if it ignores the background exposures that potentially contribute to a baseline level of DIDP in the human body. EPA's reliance on existing statutes outside of TSCA to manage exposure pathways for the general population and potentially exposed or susceptible subpopulations will underestimate risk and is scientifically unsupported.

In the preamble to the 2024 final risk evaluation framework rule, EPA acknowledged the importance of background exposures, and that these exposures can be incorporated in TSCA risk evaluations:

> it may be appropriate to consider potential background exposures from non-TSCA uses that are not within the scope of the risk evaluation as part of an aggregate exposure assessment. Likewise, EPA could consider the disproportionate impacts that background exposures may have on overburdened communities to inform the final unreasonable risk determination.[159]

EPA routinely considers exposures from products or sources that it does not regulate in assessments. For example, in its assessment and regulation of the pesticide fumigant sulfuryl fluoride, EPA's Office of Pesticide Programs ("OPP") considered all sources of exposure to fluoride, including ones EPA does not regulate (such as toothpaste). Considering these exposures was critical for accurate risk calculation and decision making—OPP proposed to terminate pesticidal uses of sulfuryl fluoride because children's total exposure to fluoride (mainly from drinking water and toothpaste) exceeded the risk cup of acceptable exposure levels.[160] EPA's plan to exclude from consideration uses of DIDP subject to statutes such as the Federal Food Drug and Cosmetics Act ignores the reality of human exposure and violates TSCA.

---

[155] 15 U.S.C. §2602 (4).
[156] 15 U.S.C. § 2625(h).
[157] NRC, Science and Decisions at 135, 136, 214, note 38 supra.
[158] Id. at 130.
[159] U.S. EPA (2024). Procedures for Chemical Risk Evaluation Under the Toxic Substances Control Act.
[160] Sulfuryl Fluoride; Proposed Order Granting Objections to Tolerances and Denying Request for Stay, 76 Fed. Reg. 3,422-01 (Jan. 19, 2011).

Thus, EPA must revise the DIDP Draft Risk Evaluation so it addresses all sources and pathways of DIDP exposure, including background exposures. TSCA, with its specific charge to consider potentially exposed or susceptible subpopulations, has a critical role to play in the protection of the general public and more susceptible groups such as infants and toddlers that are facing DIDP exposure. As we have previously detailed, established scientific principles for exposure assessment require that all known pathways of exposures be included in the assessment, or exposure will not be accurately quantified, and risk will be underestimated, particularly to potentially exposed or susceptible subpopulations.[161]

II. Comments on EPA's DINP Draft Hazard Assessment

1. EPA's non-cancer dose-response assessment for DINP is not consistent with the best available science.

a. EPA improperly excluded human epidemiology studies from dose-response assessment.

In the DINP Draft Hazard Assessment, EPA identified (primarily through public submissions to the docket) more than 50 recent human epidemiology studies of DINP non-cancer effects, using biomonitoring of urinary metabolites as measures of exposure. EPA excluded all of these studies from dose-response analysis, without any consideration of the strengths and weaknesses of each individual study:

> The Agency did not use epidemiology studies quantitatively for dose-response assessment, **primarily due to uncertainty associated with exposure characterization.** Primary sources of uncertainty include the source(s) of exposure; timing of exposure assessment that may not be reflective of exposure during outcome measurements; and use of spot-urine samples, which due to rapid elimination kinetics may not be representative of average urinary concentrations that are collected over a longer term or calculated using pooled samples. Additional uncertainty results from co-exposure to mixtures of multiple phthalates that may confound results for the majority of epidemiologic studies, which examine one phthalate and one exposure period at a time such that they are treated as if they occur in isolation.[162] (emphasis added)

EPA's blanket exclusion of an entire category of studies is scientifically inappropriate and violates the TSCA requirement to use the best available science.[163] The preamble to EPA's

[161] US EPA. (2019). Draft Toxic Substances Control Act (TSCA) Risk Evaluations and TSCA Science Advisory Committee on Chemicals (SACC); Cyclic Aliphatic Bromide Cluster (HBCD) and 1,4 Dioxane; Notice of Availability and Public Meetings. Comment submitted by Swati Rayasam, Science Associate, Program on Reproductive Health and the Environment, Department of Obstetrics, Gynecology and Reproductive Sciences, University of California, San Francisco et al. Available: https://www.regulations.gov/document?D=EPA-HQ-OPPT-2019-0237-0059 and https://www.regulations.gov/document?D=EPA-HQ-OPPT-2019-0238-0056.
[162] U.S. EPA (2024). Draft Non-cancer Human Health Hazard Assessment for Diisononyl Phthalate (DINP), pp. 12-13.
[163] 15 USC §2625(h).

recent final framework rule for conducting risk evaluations re-stated EPA's commitment to systematic review:

> EPA believes that integrating appropriate and applicable systematic review methods into the TSCA risk evaluations is critical to meeting the scientific standards as described in TSCA section 26(h) and (i)…. The principles of systematic review are well-established and include "transparent and explicitly documented methods, consistent and critical evaluation of all relevant literature, application of a standardized approach for grading the strength of evidence, and clear and consistent summative language" (Ref. 26). EPA has finalized the requirement to use and document systematic review methods to assess reasonably available information.[164]

EPA's broad exclusion of DINP epidemiology studies from dose-response analysis is contrary to the framework rule preamble and disregards the structured, consistent systematic review process that is required to evaluate the quality of relevant epidemiological studies according to pre-specified criteria. EPA has effectively ignored its systematic review process and excluded studies from dose-response assessment with an argument that demonstrates a bias against environmental epidemiology, rather than a thoughtful approach to evidence evaluation consistent with best practices in systematic review.

EPA individually assessed the merits of 53 epidemiology studies of DINP, published from 2018 to 2021, applying a pre-specified set of study quality domains and metrics that closely mirrors the approach used by EPA's IRIS program, which has been favorably reviewed by the National Academies of Sciences, Engineering, and Medicine (NASEM). EPA's overall quality determination was "Medium" or "High" for 46 of these epidemiology studies.[165] Each study was individually assessed for its exposure measurement methods (Domain 2) and treatment of potential confounding (Domain 4), and thus the issues that EPA raises in an attempt to disqualify the entire set of epidemiology studies have already been accounted for in a systematic manner using pre-specified procedures. EPA's own overall quality determinations indicate that these studies are suitable for use.

Moreover, EPA's explanation considers only alleged limitations of the DINP epidemiologic studies as a class, without considering strengths of these studies (e.g., they are conducted in humans rather than laboratory animals, at exposure levels routinely experienced by humans) or mitigating considerations (e.g. regression models that control for co-exposures; implications of exposure misclassification) that apply to the limitations. For example, the use of spot-urine samples is a limitation that is expected to result in some degree of exposure misclassification, but to the extent this occurs, it is likely to result in underestimation of risks. In general, the uncertainties in exposure characterization may result in exposure misclassification that biases dose-response estimates towards the null, but that does not mean the studies are not useful or informative and potentially strong candidates for determination of the point of departure (POD).

---

[164] U.S. EPA (2024). Procedures for Chemical Risk Evaluation Under the Toxic Substances Control Act (TSCA). 89 FR 37028.

[165] U.S. EPA (2024). Draft Risk Evaluation for Diisononyl Phthalate (DINP) – Systematic Review Supplemental File: Data Quality Evaluation Information for Human Health Hazard Epidemiology.

In an attempt to support its decision to disregard epidemiological studies, EPA cites similar decisions made in previous DINP assessments conducted by other agencies. However, the most recent of these previous assessments considered literature published only up to January 2018, whereas the 53 epidemiology studies assessed for study quality by EPA were all published from 2018-2021, and were therefore not considered in the previous assessments referenced by EPA.

EPA does not provide any scientific justification for disregarding its own conclusions regarding these studies when evaluated individually, and by overriding the findings of the systematic review process, EPA therefore violated TSCA's requirement to use the best available science.[166] EPA cannot broadly exclude epidemiologic studies from dose-response assessment in the DINP Draft Hazard Assessment, and must consider each relevant study on an individual basis as a candidate for POD derivation.

b. EPA failed to apply benchmark dose modeling to derive chronic non-cancer points of departure for risk characterization.

EPA violated its own commitment to use EPA guidance in conducting risk evaluations by not applying benchmark dose modeling to derive chronic non-cancer points of departure for risk characterization. EPA has therefore not used the best available science and leaves uncertainty regarding whether the most sensitive studies and endpoints were selected for use in estimating risks.

For risk characterization of chronic exposure to DINP, EPA proposed to use the chronic no-observed-adverse-effect level (NOAEL) for liver toxicity of 15 mg/kg-day (applied dose) from a 2-year dietary study in rats by Lington *et al.* After application of default allometric scaling, the POD is a human equivalent dose (HED) of 3.5 mg/kg-day. EPA says it has "robust overall confidence"[167] in this POD, but EPA's flawed dose-response assessment procedures for DINP do not support that conclusion.

In Table 4-5 of the DINP Draft Hazard Assessment, EPA displays 12 NOAEL and lowest-observed adverse-effect level (LOAEL) values for liver, kidney and developmental toxicity that were candidates for the chronic POD. The NOAEL HED values range from 3.5 to 48.5 mg/kg-day, and the LOAEL HEDs are from 14 to 89 mg/kg-day.[168] EPA chose the Lington *et al.* developmental toxicity NOAEL as the POD because it was more sensitive (i.e., lower) than all other candidate NOAELs and LOAELs.

Using a NOAEL as the POD rather than a benchmark dose (BMD) is not consistent with the best available science, as stated in EPA guidance[169] and reports from the NASEM.[170] By

[166] 15 USC §2625(h).
[167] U.S. EPA (2024). Draft Non-cancer Human Health Hazard Assessment for Diisononyl Phthalate (DINP), p. 89.
[168] U.S. EPA (2024). Draft Non-cancer Human Health Hazard Assessment for Diisononyl Phthalate (DINP), Table 4-5.
[169] U.S. EPA (2012). Benchmark Dose Technical Guidance.
[170] NASEM (2017). Application of systematic review methods in an overall strategy for evaluating low-dose toxicity from endocrine active chemicals, p. 158; National Research Council (2009). Science and Decisions: Advancing Risk Assessment, p. 129.

disregarding its own 2012 *Benchmark Dose Technical Guidance* in conducting dose-response analysis for DINP, EPA has violated its recent final rule *Procedures for Chemical Risk Evaluation Under the Toxic Substances Control Act (TSCA)*, which states:

> EPA will use applicable EPA guidance when conducting risk evaluations, as appropriate and where it represents the best available science.[171]

The Benchmark Dose Technical Guidance document is appropriate and represents the best available science, and it clearly states that NOAELs and LOAELs are significantly limited:

> The NOAEL is actually of little practical utility in describing toxicological dose-response relationships; it does not represent a biological threshold and cannot establish that lower exposure levels are necessarily without risk. Specific limitations of the NOAEL/LOAEL approach are well known and have been discussed extensively (Crump 1984; Gaylor 1983; Kimmel and Gaylor 1988; Leisenring and Ryan 1992; U.S. EPA 1995a):
>
> - The NOAEL/LOAEL is highly dependent on sample size. The ability of a bioassay to distinguish a treatment response from a control response decreases as sample size decreases, so the NOAEL for a compound (and thus the POD, when based on a NOAEL) will tend to be higher in studies with smaller numbers of animals per dose group.
> - More generally, the NOAEL/LOAEL approach does not account for the variability and uncertainty in the experimental results that are due to characteristics of the study design such as dose selection, dose spacing, and sample size.
> - NOAELs/LOAELs do not correspond to consistent response levels for comparisons across studies/chemicals/endpoints, and the observed response level at the NOAEL or LOAEL is not considered in the derivation of RfDs/RfCs.
> - Other dose-response information from the experiment, such as the shape of the dose-response curve (e.g., how steep or shallow the slope is at the BMD, providing some indication of how near the POD might be to an inferred threshold), is not taken into account…
> - While the NOAEL has typically been interpreted as a threshold (no-effect level), simulation studies (e.g., Leisenring and Ryan 1992; study designs involving 10, 20, or 50 replicates per dose group) and re-analyses of developmental toxicity bioassay data (Gaylor 1992; Allen et al. 1994a; studies involving approximately 20 litters per dose group) have demonstrated that the rate of response above control at doses fitting the criteria for NOAELs, for a range of study designs, is about 5–20% on average, not 0%.[172]

The Benchmark Dose Technical Guidance further states that use of a BMD/BMDL as a POD is preferred and a NOAEL or LOAEL should be considered as a POD only if BMD modeling is conducted and is unable to produce a BMD estimate, and requires justification:

---

[171] 40 CFR § 702.37.
[172] U.S. EPA (2012). Benchmark Dose Technical Guidance, p. 4.

> Because of the limitations of the NOAEL/LOAEL approach discussed earlier, the BMD approach is preferred to the NOAEL/LOAEL approach… there are some instances in which reliable BMDs cannot be estimated and the NOAEL/LOAEL approach might be warranted…In such cases, the NOAEL/LOAEL approach might be used, while recognizing its limitations and the limitations of the dataset.[173]
>
> Resorting to the NOAEL/LOAEL approach does not resolve a data set's inherent limitations, but it conveys that there are limitations with the data set.[174]
>
> At times, modeling will not yield useful results and the NOAEL/LOAEL approach might be considered, although the data gaps and inherent limitations of that approach should be acknowledged.[175]
>
> In some cases, modeling attempts may not yield useful results. When this occurs and the most biologically relevant effect is from a study considered adequate but not amenable to modeling, the NOAEL (or LOAEL) could be used as the POD. The modeling issues that arose should be discussed in the assessment, along with the impacts of any related data limitations on the results from the alternate NOAEL/LOAEL approach.[176]

EPA cited the Benchmark Dose Technical Guidance in a previous TSCA risk evaluation to describe the preference for a BMD over a NOAEL:

> As outlined in EPA guidance, the BMD approach overcomes many of the limitations inherently associated with the NOAEL/LOAEL approach, and thus is the preferred method for establishing a POD for use in risk assessment.[177]

EPA's 2022 handbook for conducting chemical hazard assessments for the Integrated Risk Information System (IRIS) reinforces these key points:

> As discussed in detail in Section 1.2 of EPA's Benchmark Dose Technical Guidance (U.S. EPA, 2012b), dose-response modeling (i.e., benchmark dose modeling) is the preferred approach for deriving points of departures given several limitations in the no-observed adverse-effect level/ lowest-observed-adverse-effect level (NOAEL/LOAEL) approach.[178]
>
> Basis of the POD: A modeled BMDL is preferred over a NOAEL, which is in turn preferred over a LOAEL.[179]

Reports from the NASEM also state the advantages of BMD modeling. The NASEM report on low-dose toxicity of endocrine active chemicals (which was the source of the BMDL selected as

[173] U.S. EPA (2012). Benchmark Dose Technical Guidance, p. 6.
[174] U.S. EPA (2012). Benchmark Dose Technical Guidance, p. 12.
[175] U.S. EPA (2012). Benchmark Dose Technical Guidance, p. 30.
[176] U.S. EPA (2012). Benchmark Dose Technical Guidance, p. 40.
[177] U.S.EPA (2020). Risk Evaluation for n-Methylpyrrolidone (2-Pyrrolidinone, 1-Methyl-) (NMP), p. 262.
[178] U.S. EPA (2022). ORD Staff Handbook for Developing IRIS Assessments, p. 8-1.
[179] U.S. EPA (2022). ORD Staff Handbook for Developing IRIS Assessments, p. 8-18.

POD for acute effects of DINP) discusses the deficiencies of the NOAEL/LOAEL approach for risk estimation:

> The use of LOAELs and NOAELs is less than ideal because they depend highly on individual study-design characteristics; therefore, apparent differences among studies might be explained by design differences, such as sample size or dose spacing, rather than true inconsistency.[180]

In the 2009 report *Science and Decisions*, the National Academies highlighted the adoption of the BMD approach as an important improvement in risk assessment methodology:

> Another refinement in dose-response assessment has been the derivation of the RfD or low-dose cancer risk from a POD that is calculated using BMD methodology (EPA 2000a). In noncancer risk assessment, this approach has the advantage of making better use of the dose-response evidence available from bioassays than do calculations based on NOAELs. It also provides additional quantitative insight into the risk presented in the bioassay at the POD because for quantal end points the POD is defined in terms of a given risk for the animals in the study.[181]

EPA did conduct BMD modeling of multiple endpoints from the Lington *et al*. study, but then chose to use the NOAEL rather than a lower bound BMD value (BMDL) as the POD, which directly conflicts with the EPA Benchmark Dose Technical Guidance, EPA IRIS handbook, previous TSCA risk evaluations, and NASEM recommendations. The EPA-estimated BMDLs for some endpoints were lower than the NOAEL. EPA chose not to use the BMDL of 8.6 mg/kg-day (applied dose) for spongiosis hepatis in the liver or the BMDL of 15.5 mg/kg-day for serum ALT at 6-month sacrifice[182] as the chronic POD, instead using the NOAEL as the chronic POD without appropriate scientific justification:

> The wide variability in BMDLs and uncertainty in several modelled outcomes (*i.e.,* BMD/BMDL ratios greater than 3) reduce EPA's confidence in using the BMD modeling results for establishing a POD, and further affirm the use of the NOAEL for establishing the POD.[183]

Variability in BMDLs across endpoints is not a valid justification for using a NOAEL rather than a BMDL; EPA guidance (see above) instead emphasizes the strong preference for using a BMDL rather than a NOAEL. EPA's mention of BMD/BMDL ratios is not supported by EPA guidance, and furthermore is not valid because the BMD/BMDL ratio for increased serum ALT is only 1.5 (23.4 / 15.5 = 1.5).[184]

---

[180] NASEM (2017). Application of systematic review methods in an overall strategy for evaluating low-dose toxicity from endocrine active chemicals, p. 158.
[181] National Research Council (2009). Science and Decisions: Advancing Risk Assessment, p. 129.
[182] U.S. EPA (2024). Draft Non-cancer Human Health Hazard Assessment for Diisononyl Phthalate (DINP), Table_Apx E-1.
[183] U.S. EPA (2024). Draft Non-cancer Human Health Hazard Assessment for Diisononyl Phthalate (DINP), p. 89.
[184] U.S. EPA (2024). Draft Non-cancer Human Health Hazard Assessment for Diisononyl Phthalate (DINP), Table_Apx E-7.

EPA's attempt to justify using a NOAEL as POD continues:

> EPA considers it more appropriate to use the NOAEL of 15 mg/kg-day instead of the BMD05 of 12 mg/kg-day because the NOAEL supports the suite of effects on the liver occurring at 152 mg/kg-day instead of being based on the single effect of spongiosis hepatis with its associated uncertainty regarding human relevance.[185]

This explanation also is not scientifically valid. First, EPA uses an inappropriate comparison of the NOAEL to a BMD value, when the lower-bound estimate of the BMD (i.e. the $BMDL_{05}$) is the appropriate choice. Second, the mention of the term "suite of effects" disregards the fact that the BMD analysis shows that, as one would expect, some liver effects are more sensitive than others; the use of the term "suite of effects" averages over multiple outcomes to obscure the most sensitive outcomes, contrary to the objective of selecting the most sensitive endpoint. Finally, EPA does not give any rationale for disregarding the increased serum ALT BMDL of 15 mg/kg-day in selecting the POD.

Further, EPA did not conduct BMD modeling for any of the candidate chronic PODs other than the liver effects from Lington *et al.* This means that EPA has not applied the best available science to determine the most sensitive endpoint, as it selected the POD without conducting appropriate dose-response analysis and instead relied on comparisons of NOAELs and LOAELs. EPA's Benchmark Dose Technical Guidance clearly states that identification of the most sensitive endpoint cannot be based on comparisons of NOAELs, and that all candidate values should be evaluated based on BMD modeling:

> The apparent relative sensitivities of endpoints based on NOAELs/LOAELs may not correspond to the same relative sensitivities based on BMDs or BMDLs after BMD modeling; therefore, relative sensitivities of endpoints cannot necessarily be judged a priori. For example, differences in slope (at the BMR) among endpoints could affect the relative values of the BMDLs. Selected endpoints from different studies that have the potential to be used in the determination of a POD(s) should all be modeled.[186]

A BMDL is frequently lower than the NOAEL for the same endpoint, and frequently much lower than the LOAEL for the same endpoint. Without BMD modeling, EPA is unable to make a scientific determination of whether the findings from Lington *et al*. study are more sensitive than the liver effects from Bio/dynamics 1987 (NOAEL HED = 6.4 mg/kg-day), kidney lesions from Hazleton labs (*LOAEL* HED = 14.2 mg/kg-day), developmental effects from Waterman *et al*. (*LOAEL* HED = 31.4 mg/kg-day), or other candidate endpoints.[187] The scientifically appropriate method for selecting the POD based on the most sensitive endpoint would be to estimate a BMDL for each endpoint, and then select the lowest value.

[185] U.S. EPA (2024). Draft Non-cancer Human Health Hazard Assessment for Diisononyl Phthalate (DINP), p. 89.
[186] U.S. EPA (2012). Benchmark Dose Technical Guidance, p. 15.
[187] U.S. EPA (2024). Draft Non-cancer Human Health Hazard Assessment for Diisononyl Phthalate (DINP), Table 4-5.

By disregarding existing EPA guidance and NASEM recommendations that state BMD modeling is the most scientifically appropriate approach for determining the POD, EPA violates the TSCA section 26(h) scientific standards which direct that the Agency:

> Shall use scientific information, technical procedures, measures, methods, protocols, methodologies, or models, employed in a manner consistent with the best available science.[188]

Further, EPA's recently promulgated revisions to the framework rule for TSCA risk evaluations states that:

> EPA will document that the risk evaluation is consistent with the best available science.[189]

EPA cannot ensure that the final DINP Draft Hazard Assessment meets this requirement unless it has implemented BMD modeling in the process of selecting a POD.

c. EPA should apply best available methods to generate quantitative estimates of non-cancer risks for varying levels of exposure to DINP.

In its TSCA risk evaluations, EPA typically calculates a margin of exposure (MOE) for each condition of use (COU). The MOE is calculated as:

> Margin of Exposure = Non-cancer point of departure / Human exposure.

The MOE approach is a scientifically deficient method for characterizing risk and is inconsistent with TSCA's requirements to use the "best available science"[190] and to ensure protection of "potentially exposed and susceptible subpopulations" ("PESS").[191]

Use of the MOE, which relies on a POD with no extrapolation to lower doses, is a simplistic approach that only compares the POD to the exposure level and judges whether this ratio is interpreted as a human health risk of concern" or if "risk is not considered to be of concern and mitigation is not needed."[192] The MOE does not estimate the proportion of the exposed population projected to experience a specified health endpoint or the number of individuals affected, and it perpetuates the scientifically flawed notion that a "safe" or "no risk" level of chemical exposure can be identified for a diverse exposed population.[193,194]

---

[188] 15 U.S.C. § 2625(h).
[189] U.S. EPA (2024). Procedures for Chemical Risk Evaluation Under the Toxic Substances Control Act (TSCA). 89 FR 37028, May 3, 2024, § 702.37(a)(2).
[190] 15 U.S.C. §2625 (h).
[191] 15 U.S.C. §2602 (12).
[192] U.S. EPA (2024). Draft Risk Evaluation for Diisodecyl Phthalate (DIDP), p. 113.
[193] Woodruff, T. J., Rayasam, S. D. G., Axelrad, D. A., Koman, P. D., Chartres, N., et al.. (2023). A science-based agenda for health-protective chemical assessments and decisions: overview and consensus statement. Environ Health, 21(Suppl 1), 132. https://doi.org/10.1186/s12940-022-00930-3.
[194] McGartland, A., Revesz, R., Axelrad, D. A., Dockins, C., Sutton, P., Woodruff, T. J. (2017). Estimating the health benefits of environmental regulations. Science, 357(6350), 457-458. https://doi.org/10.1126/science.aam8204.

The National Academies[195] and the World Health Organization[196] ("WHO") have outlined more robust methods for risk estimation that more accurately account for variability and vulnerability across the human population and have been demonstrated in published case studies.[197,198 ,199,200] We applied the WHO methodology to the DINP liver toxicity endpoints of spongiosis hepatis (a type of liver lesion) and increased serum ALT (a biomarker indicating liver damage), using the BMD and BMDL values reported by EPA, to estimate risk-specific doses for several levels of incidence (e.g. 1%, 0.1%, etc.).

Our analysis (see Technical Appendix B for details; all reported doses are HEDs) found that:

1. 0.44 mg/kg-day is the lower bound (95% confidence) chronic human dose at which liver lesions are expected in 5% of the exposed population, and 0.17 mg/kg-day is the lower bound (95% confidence) chronic human dose at which reduced serum ALT is expected in 5% of the exposed population;
2. 0.18 mg/kg-day is the lower bound (95% confidence) chronic human dose at which liver lesions are expected in 1% of the exposed population, and 0.065 mg/kg-day is the lower bound (95% confidence) chronic human dose at which reduced serum ALT is expected in 1% of the exposed population;
3. 0.12 mg/kg-day is the lower bound (95% confidence) chronic human dose at which liver lesions are expected in 0.5% of the exposed population, and 0.04 mg/kg-day is the lower bound (95% confidence) chronic human dose at which reduced serum ALT is expected in 0.5% of the exposed population;
4. 0.06 mg/kg-day is the lower bound (95% confidence) chronic human dose at which liver lesions are expected in 0.1% of the exposed population, and 0.02 mg/kg-day is the lower bound (95% confidence) chronic human dose at which reduced serum ALT is expected in 0.1% of the exposed population;
5. 0.02 mg/kg-day is the lower bound (95% confidence) chronic human dose at which liver lesions are expected in 0.01% (1-in-10,000) of the exposed population, and 0.008 mg/kg-day is the lower bound (95% confidence) chronic human dose at which reduced serum ALT is expected in 0.01% of the exposed population;
6. 0.01 mg/kg-day is the lower bound (95% confidence) chronic human dose at which liver lesions are expected in 0.001% (1-in-100,000) of the exposed population, and 0.003

[195] National Research Council (2009). Science and Decisions: Advancing Risk Assessment, Chapter 5.
[196] WHO. (2017). Guidance document on evaluating and expressing uncertainty in hazard characterization. Harmonization project document 11, 2nd edition. https://www.who.int/publications/i/item/9789241513548.
[197] Chiu WA, Axelrad DA, Dalaijamts C, Dockins C, Shao K, Shapiro AJ, Paoli G. Beyond the RfD: broad application of a probabilistic approach to improve chemical dose-response assessment for non-cancer effects. Environmental Health Perspectives, 2018 June;126(6):067009. doi:10.1289/EHP3368.
[198] Nielsen, G. H., Heiger-Bernays, W. J., Levy, J. I., White, R. F., Axelrad, D. A., Lam, J., Chartres, N., Abrahamsson, D. P., Rayasam, S. D. G., Shaffer, R. M., Zeise, L., Woodruff, T. J., Ginsberg, G. L. (2023). Application of probabilistic methods to address variability and uncertainty in estimating risks for non-cancer health effects. Environ Health, 21(Suppl 1), 129. https://doi.org/10.1186/s12940-022-00918-z.
[199] Blessinger, T., Davis, A., Chiu, W. A., Stanek, J., Woodall, G. M., Gift, J., Thayer, K. A., Bussard, D. (2020). Application of a unified probabilistic framework to the dose-response assessment of acrolein. Environ Int, 143,105953. https://doi.org/10.1016/j.envint.2020.105953.
[200] Ginsberg, G. L. (2012). Cadmium risk assessment in relation to background risk of chronic kidney disease. J Toxicol Environ Health A, 75(7),374-390.

> mg/kg-day is the lower bound (95% confidence) chronic human dose at which reduced serum ALT is expected in 0.001% of the exposed population.

The implications of these risk values can be understood by comparison with the exposure levels considered by EPA to represent negligible risk. EPA's assessment uses a POD of 3.5 mg/kg-day (HED) and a benchmark MOE of 30,[201] meaning that EPA concludes "risk is not considered to be of concern and mitigation is not needed"[202] for any exposure below 0.12 mg/kg-day (3.5 mg/kg-day / 30 = 0.12 mg/kg-day). Our analysis finds that an exposure of 0.12 mg/kg-day is equal to the lower-bound dose for the 0.5% (1-in-200) risk level for spongiosis hepatis lesions, and an exposure of 0.12 mg/kg-day is greater than the lower-bound dose for the 1% (1-in-100) risk level increased serum ALT. These risks far exceed EPA's usual target range of protection for carcinogenic risks of 1-in-10,000 to 1-in-1,000,000.[203]

EPA should apply the WHO framework to the DINP liver endpoints from the Lington *et al*. study. EPA should also conduct BMD modeling for other DINP candidate studies, and use the BMD outputs to apply the WHO framework to other non-cancer endpoints of DINP.

2. EPA did not apply the best available science to identify and evaluate relevant and useful health effects studies for DINP.

a. EPA has not released a systematic review protocol for DINP. This means that EPA has employed methods in preparing the DINP hazard assessment that have not been disclosed to the public or to the SACC.

EPA says that its procedures for identifying and reviewing the non-cancer effects evidence for DINP are described in a systematic review protocol:

> EPA's process for considering and incorporating new DINP literature is described in the *Draft Risk Evaluation for Diisononyl Phthalate (DINP) – Systematic Review Protocol* (also referred to as the Draft DINP Systematic Review Protocol).[204]

A systematic review protocol is absolutely necessary for the process of conducting a TSCA risk evaluation, so it is appropriate for the draft non-cancer hazard assessment to reference a protocol. However, EPA has not released the cited protocol to the public for the current comment period, even though it has released many other supplemental files in the docket. It is unclear why EPA has withheld the protocol, or why the hazard assessment document cites a protocol that is not available.

---

[201] U.S. EPA (2024). Draft Non-cancer Human Health Hazard Assessment for Diisononyl Phthalate (DINP), Table ES-1.
[202] U.S. EPA (2024). Draft Risk Evaluation for Diisodecyl Phthalate (DIDP), p. 113.
[203] U.S. EPA (2024). Unreasonable Risk Determination of the Draft Risk Evaluation for Formaldehyde, p. 13.
[204] U.S. EPA (2024). Draft Non-cancer Human Health Hazard Assessment for Diisononyl Phthalate (DINP), p. 11.

The Institute of Medicine's list of best practices for systematic review include making a protocol available for public comment before conducting the review, and making the final protocol publicly available:

> Provide a public comment period for the protocol and publicly report on disposition of comments.
>
> Make the final protocol publicly available, and add any amendments to the protocol in a timely fashion.[205]

Without a protocol available, EPA has failed to be transparent regarding the methods applied in preparing the draft non-cancer and cancer hazard assessments for DINP. The information on methods that is provided in the DINP Draft Hazard Assessment is a very limited and unclear summary that cannot be considered a substitute for a protocol, and includes some concerning elements. For example, Figure 1-1seems to indicate that study quality evaluation is conducted only after substantial narrowing of the evidence base to focus on selected endpoints and studies used for dose-response assessment.[206] In addition, the subsequent text indicates that EPA included studies for hazard assessment without assessing the quality of those studies:

> EPA did not conduct data quality evaluations for studies…not considered sensitive for subsequent POD selection. However, these studies were still reviewed and integrated into the hazard identification process.[207]

EPA did release a draft systematic review protocol for diisodecyl phthalate (DIDP) at the same time that it released the DINP Draft Hazard Assessment. Although the DIDP protocol is severely flawed, it likely indicates methods used for the DINP Draft Hazard Assessment that are not described in the available draft DINP documents. Assuming that the same methods described in the DIDP protocol were applied to DINP, there are several critical methodological steps to the DINP Draft Hazard Assessment that have not been disclosed to the public or to the SACC. These include:

- Application of unclear "further filtering" procedures to exclude certain health effects studies that EPA had determined were relevant, with no explanation for why further filtering was necessary;
- Exclusion of relevant health effects studies from the assessment if they did not provide data that EPA considered useful for dose-response assessment;
- Exclusion of studies assigned an overall rating of "uninformative," even though EPA does not describe how overall ratings are determined, and even though this term has been assigned to studies in previous TSCA risk evaluations (e.g. formaldehyde) where EPA actually did find useful for hazard and dose-response assessment.

[205] Institute of Medicine (2011). Finding what works in health care: Standards for systematic reviews, p. 75.
[206] U.S. EPA (2024). Draft Non-cancer Human Health Hazard Assessment for Diisononyl Phthalate (DINP), p. 17.
[207] U.S. EPA (2024). Draft Non-cancer Human Health Hazard Assessment for Diisononyl Phthalate (DINP), p. 17.

In addition, for DIDP, only the protocol provides a general overview of the process by which included studies were identified. For DINP, the non-cancer hazard document is unclear on this point. For example, EPA reports that it identified 25 studies of non-cancer liver toxicity for DINP,[208] but failed to disclose how this set of studies was assembled – e.g., it is unclear how many studies were identified through review of assessments by other agencies, and how many were identified from EPA's search of studies published from 2014-2019. This is a critical lack of transparency in EPA's assessment and is inconsistent with best scientific practices. Although significantly flawed, Figure 4-6 in the DIDP protocol did attempt to indicate how its body of toxicology studies was assembled – information which is not available for DINP but is presumably included in the unreleased DINP protocol.

An additional public comment period and peer review of the DINP Draft Hazard Assessment should be conducted after EPA has released the systematic review protocol.

For future TSCA risk evaluations, EPA must publish a chemical-specific systematic review protocol for public comment *before* completing the draft risk evaluation, as recommended by the Institute of Medicine and the NASEM as a best practice for systematic review.[209,210]

The TSCA program should follow the established procedures of EPA's IRIS program, which makes a draft protocol for each assessment publicly available in advance of its release for public comment. Following the public comment process, the IRIS program then publishes an updated protocol, as needed. For example, for the IRIS assessments of five per- and polyfluoroalkyl substances ("PFAS"), a draft protocol was made available for public comment for 45 days. The IRIS program then followed up with a revised protocol to address public comments, with documentation of the changes, that was published before the release of the PFAS draft assessments.[211] EPA should be following this same approach for all TSCA risk evaluations.

b. EPA did not conduct a comprehensive and up-to-date literature search.

The need for transparent, consistent and comprehensive approaches to identifying health effects literature has been a key driver for increased adoption of systematic review methods in environmental health assessments over the past 15 years.[212,213,214] EPA's assessment of DINP is a concerning step backwards in this area, as the approach to identifying evidence is not clear,

---

[208] U.S. EPA (2024). Draft Non-cancer Human Health Hazard Assessment for Diisononyl Phthalate (DINP), p. 42.
[209] Institute of Medicine (2011). Finding what works in health care: Standards for systematic reviews.
[210] National Research Council (2014). Review of EPA's Integrated Risk Information System (IRIS) process.
[211] U.S. EPA (2021). Systematic Review Protocol for the PFAS IRIS Assessments. https://cfpub.epa.gov/ncea/iris_drafts/recordisplay.cfm?deid=345065 (accessed 1 February 2024).
[212] National Research Council (2011). Review of the Environmental Protection Agency's Draft IRIS Assessment of Formaldehyde.
[213] Woodruff TJ, Sutton P; Navigation Guide Work Group. An evidence-based medicine methodology to bridge the gap between clinical and environmental health sciences. Health Affairs 2011 May;30(5):931-7. doi: 10.1377/hlthaff.2010.1219.
[214] Rooney AA, Boyles AL, Wolfe MS, Bucher JR, Thayer KA. 2014. Systematic review and evidence integration for literature-based environmental health science assessments. Environ Health Perspect 122:711–718.

consistent or comprehensive. Based on the inconsistent procedures applied, it is unlikely that EPA would have identified and included all relevant health effects studies. This indicates critical deficiencies in the EPA systematic review protocol and the DINP Draft Hazard Assessment.

For the DINP Draft Hazard Assessment, EPA relied on non-EPA assessments of DINP completed in 2018 or earlier, and a literature search that was conducted in 2019 and has not been updated since.

For identifying epidemiological studies, EPA described its procedures as follows:

> To identify and integrate human epidemiologic data into the draft DINP Risk Evaluation, EPA first reviewed existing assessments of DINP conducted by regulatory and authoritative agencies… **most** of these assessments have been subjected to peer-review and/**or** public comment periods and have employed formal systematic review protocols.[215] (emphasis added)

> Next, EPA sought to identify new population, exposure, comparator, and outcome (PECO)-relevant literature published since the most recent existing assessment(s) of DINP by applying a literature inclusion cutoff date. For DINP, the applied cutoff date was based on existing assessments of epidemiologic studies of phthalates by Health Canada (2018a, b), which included literature up to January 2018….New PECO-relevant literature published between 2018 to 2019 was identified through the literature search conducted by EPA in 2019, as well as references published between 2018 to 2023 that were submitted with public comments to the DINP Docket…were evaluated for data quality.[216]

EPA therefore conducted a comprehensive literature search only for studies published in a time period of less than 2 years. As a result, the set of epidemiology studies consists of three inconsistent subsets:

- Studies published prior to January 2018 – are included in EPA's assessment only if they were included in the assessments conducted by other agencies. The assessments used by EPA to identify studies were not necessarily peer-reviewed and were not necessarily systematic reviews. EPA did not assess the quality of the studies identified by these other assessments. EPA did not consider any studies published before 2018 if they were not discovered by or not included in previous assessments for any reason.
- Studies published from January 2018 – 2019 – EPA conducted its own search of the literature and applied its own inclusion/exclusion criteria.
- Studies published from 2019 to May 2024 – are included in EPA's assessment only if they were submitted to the EPA docket.

[215] U.S. EPA (2024). Draft Non-cancer Human Health Hazard Assessment for Diisononyl Phthalate (DINP), p. 11.
[216] U.S. EPA (2024). Draft Non-cancer Human Health Hazard Assessment for Diisononyl Phthalate (DINP), p. 12.

Thus, only those epidemiology studies published in a span of up to 24 months were identified and evaluated through a comprehensive process following an EPA protocol (in this case, a protocol that is not yet available). For earlier studies (before 2018), EPA relied entirely on the Health Canada and other agency assessments and did not apply its own search, inclusion/exclusion and study evaluation procedures. For later studies (after some unspecified date in 2019), EPA did not conduct a search but included only those studies that were submitted by the public to EPA. This is not a clear, comprehensive or consistent approach to identifying the epidemiological evidence relevant to assessing the health effects of DINP. A further concern is that these inconsistent procedures for identifying epidemiological evidence were ultimately relevant only to the identification of DINP hazards, since EPA subsequently excluded **all** epidemiological studies from consideration for dose-response assessment, without consideration of the merits of individual studies (see Section 1a. above).

For identifying toxicology studies, EPA applied a similar process:

> EPA first reviewed existing assessments of DINP conducted by various regulatory and authoritative agencies…The purpose of this review was to identify sensitive and human relevant hazard outcomes associated with exposure to DINP, and identify key studies used to establish PODs for estimating human risk… **most** of these assessments have been subjected to external peer-review and/or public comment periods **but have not employed formal systematic review protocols**.[217] (emphasis added)

> EPA used the 2015 Health Canada assessment (EC/HC, 2015) as the key starting point for this draft document. The Health Canada assessment included scientific literature up to August 2014...Therefore, EPA considered literature published between 2014 to 2019 further…EPA reviewed new studies published between 2014 and 2019 and extracted key study information.[218]

EPA therefore conducted a comprehensive literature search only for studies published in a 5-year span. As a result, DINP toxicology studies are divided into three inconsistent subsets:

- Studies published up to mid-2014 – included only if they were included in the previous assessment by Health Canada. Additionally, EPA did not consider any studies published before mid-2014 if they were not discovered by or not included in the previous assessments for any reason.
- Relevant studies published from mid-2014 to 2019 – EPA conducted its own search of the literature and applied its own inclusion/exclusion criteria.
- Studies published after 2019 (date unspecified) – were not considered at all.

[217] U.S. EPA (2024). Draft Non-cancer Human Health Hazard Assessment for Diisononyl Phthalate (DINP), p. 16.
[218] U.S. EPA (2024). Draft Non-cancer Human Health Hazard Assessment for Diisononyl Phthalate (DINP), p. 17.

Thus, only those toxicology studies published in a span of approximately 5 years were identified and evaluated through a comprehensive process following an EPA protocol (in this case, a protocol that is not yet available). For earlier toxicology studies (before mid-2014), EPA relied entirely on assessments by other agencies and did not apply its own search, inclusion/exclusion and study evaluation procedures. Toxicology studies published after an unspecified date in 2019 were not included at all. This is not a clear, comprehensive or consistent approach to identifying the toxicology evidence relevant to assessing the health effects of DINP.

For both epidemiology and toxicology, studies were treated differently based only on their date of publication. In addition, the procedures for epidemiology differed significantly from those for toxicology; for example, some post-2019 epidemiology studies were included (but not necessarily all relevant studies, since a search was not conducted), whereas no post-2019 toxicology studies were included. Any toxicological findings on DINP published in the past 5 years were simply not considered by EPA, which is not consistent with the best available science; recent guidance on conducting systematic reviews in environmental health recommends that literature searches should be updated no more than 12 months before publication of a review.[219] Collectively, EPA's practices run a high risk of failing to include all relevant health effects studies and/or treating relevant studies differently in the DINP risk evaluation.

c. EPA relied on assessments conducted by other agencies to exclude studies, without supporting justification.

EPA reviewed DINP health effects assessments conducted by Canada, Australia, multi-lateral European agencies, the U.S. CPSC and the U.S. NTP as part of conducting the DINP Draft Hazard Assessment. Epidemiology studies published before 2019 and toxicology studies published before mid-2014 were included in the TSCA risk evaluation only if they were included in these previous assessments. Studies that were not identified in searches conducted in the previous assessments and studies that were excluded from the previous assessments for any reason were not considered at all by EPA.

In principle, the use of previous assessments can be a useful part of conducting a TSCA risk evaluation, but the previous assessments must be carefully evaluated against a pre-specified set of criteria to determine whether they are of sufficient quality, and the resulting risk evaluation must still employee procedures that are transparent, comprehensive, consistent and unbiased, and must meet the TSCA section 26(h) scientific standards which direct that the Agency:

> Shall use scientific information, technical procedures, measures, methods, protocols, methodologies, or models, employed in a manner consistent with the best available science.[220]

However, EPA notes that the previous assessments it used were not systematic reviews, and not all were peer reviewed. EPA also does not provide adequate justification for its use of previous

[219] P. Whaley, et al. Recommendations for the conduct of systematic reviews in toxicology and environmental health research (COSTER). Environment International 143 (2020), 105926. https://doi.org/10.1016/j.envint.2020.105926.
[220] 15 U.S.C. § 2625(h).

DINP assessments to substitute for conducting its own comprehensive systematic review to identify and evaluate health effects evidence.

The 2023 NASEM report *Building Confidence in New Evidence Streams for Human Health Risk Assessment* demonstrates an appropriate process for evaluating the quality of previous assessments. After conducting a comprehensive search for prior reviews satisfying a pre-specified PECO (population, exposure, comparator, outcome) statement, the NASEM applied AMSTAR 2 (A MeaSurement Tool to Assess systematic Reviews) to assess the methodological quality of each relevant review.[221] AMSTAR 2 was also applied by the NASEM in multiple prior reports on environmental health assessment.[222,223,224] In order to establish that it is appropriate to use previous assessments as part of a TSCA risk evaluation, EPA must apply this type of process to determine whether the previous assessments are consistent with the best available science, as required by TSCA.[225]

d. EPA used deficient inclusion and exclusion criteria for health effects evidence that inappropriately excluded important toxicity endpoints.

The DINP Draft Hazard Assessment does not provide the PECO statement that was used to identify epidemiology studies published from 2018-2019 and toxicology studies published from 2014-2019. The document states:

> EPA first screened titles and abstracts and then full texts for relevancy using PECO screening criteria described in the Draft DINP Systematic Review Protocol.[226]

However, as discussed above, EPA has not made the DINP systematic review protocol available to the public during the current public comment period. A PECO statement was provided in the broader 2021 TSCA Draft Systematic Review Protocol, which EPA has never revised to address public comments and more than 200 SACC recommendations. Since EPA has not conducted a search for DINP health evidence since 2019, we assume that the 2021 PECO was applied in preparing the draft hazard assessment.

PECO statements play a critical role in conducting a systematic review as they provide criteria for screening the literature search results to identify which studies are relevant (included in the risk evaluation) and not relevant (excluded from further consideration). The PECO statement for DINP is deficient and excludes a broad range of important toxicity outcomes from consideration in the draft risk evaluation.

---

[221] NASEM 2023). Building Confidence in New Evidence Streams for Human Health Risk Assessment: Lessons Learned from Laboratory Mammalian Toxicity Tests.
[222] NASEM (2019). Review of DOD's Approach to Deriving an Occupational Exposure Level for Trichloroethylene.
[223] NASEM (2021). The Use of Systematic Review in EPA's Toxic Substances Control Act Risk Evaluations.
[224] NASEM (2022). Guidance on PFAS Exposure, Testing, and Clinical Follow-Up.
[225] 15 U.S.C. § 2625(h).
[226] U.S. EPA (2024). Draft Non-cancer Human Health Hazard Assessment for Diisononyl Phthalate (DINP), p. 17.

The outcome component of the PECO statement for DINP health effects evidence provides the following criteria for inclusion and exclusion of studies:

> **Human:** All health outcomes (cancer and non-cancer) at the organ level or higher.
> **Animal and Plants:** All apical biological effects (effects measured at the organ level or higher)
> and bioaccumulation from laboratory studies with concurrently measured media and/or tissue concentrations. Apical endpoints include but are not limited to reproduction, survival, and growth.
> **Screener note:**
> - Measurable biological effects relevant for humans, animals and plants may include but are not limited to: mortality, behavioral, population, physiological, growth, reproduction, systemic, point of contact (irritation and sensitization) effects.
> - Effects measured at the cellular level of biological organization and below are to be tagged as supplemental, mechanistic.[227] (emphasis added)

By limiting the relevant human and animal studies to those with "apical" effects or those with effects at the "organ level or higher," EPA appears to be excluding studies of important biochemical markers and other outcomes at the cellular level that are strong indicators of hazards and which have commonly been used as critical effects in previous EPA hazard assessments, including TSCA risk evaluations (see examples below).

EPA's PECO statement provides very limited guidance for screeners on what effects are to be considered "apical" or "organ-level." The PECO says: "Apical endpoints include but are not limited to reproduction, survival, and growth" and "Measurable biological effects relevant for humans, animals and plants may include but are not limited to: mortality, behavioral, population, physiological, growth, reproduction, systemic, point of contact (irritation and sensitization) effects."[228] The 2021 TSCA Draft Systematic Review Protocol provides no further guidance on which outcomes are to be considered apical or organ-level, and which outcomes are to be considered cellular-level.

The NASEM has defined an apical end point as "An observable outcome in a whole organism, such as a clinical sign or pathologic state, that is indicative of a disease state that can result from exposure to a toxicant,"[229] and identified "tumors, birth defects, and neurologic impairments"[230] as examples. No biochemical measures or early biological changes were mentioned among the examples.

The definition of an apical effect appears to be narrower than the definition of an adverse effect provided by the EPA IRIS program: "a biochemical change, functional impairment, or

[227] U.S. EPA (2021). Draft Systematic Review Protocol Supporting TSCA Risk Evaluations for Chemical Substances, Table_Apx H-47.
[228] U.S. EPA (2021). Draft Systematic Review Protocol Supporting TSCA Risk Evaluations for Chemical Substances, Table_Apx H-47.
[229] National Research Council (2007). Toxicity Testing in the 21st Century: A Vision and a Strategy, p. 38.
[230] National Research Council (2007). Toxicity Testing in the 21st Century: A Vision and a Strategy, p. 177.

pathologic lesion that affects the performance of the whole organism, or reduces an organism's ability to response to an additional environmental challenge."[231] The definition of adverse effect includes, for example, "a biochemical change;" such effects appear to be excluded from the DINP Draft Hazard Assessment as they would likely be considered cellular-level effects rather than organ-level or apical effects

Biochemical and/or cellular-level outcomes have been identified as critical effects in numerous past EPA hazard assessments, including some of the completed TSCA risk evaluations. Examples of these outcomes and past assessments include:

- reduced male fetal testosterone or adult male testosterone levels (2018 and 2019 IRIS staff published systematic reviews of health effects of phthalates, 2023 draft approach to cumulative risk assessment of phthalates under TSCA)[232,233,234]
- reduced thyroid hormone levels (2020 TSCA risk evaluation of HBCD; 2021 toxicity assessment of PFBS) [235,236]
- decreased erythrocyte counts and hemoglobin (2020 TSCA risk evaluation of perchloroethylene)[237]
- measures of immune function, such as increases in immunoglobulin E, lymphocytes, natural killer cells, and interlukin-4 levels (2020 TSCA risk evaluation of perchloroethylene)[238]
- decreased sperm quality or concentration (2020 TSCA risk evaluations of trichloroethylene and perchloroethylene; 2018 and 2019 IRIS staff published systematic reviews of health effects of phthalates)[239,240,241,242]
- acetylcholinesterase inhibition (numerous assessments of pesticides, including cumulative risk assessments of organophosphate and carbamate pesticides)[243,244]

---

[231] U.S. EPA. IRIS Glossary. https://www.epa.gov/iris/iris-glossary.
[232] Radke EG, Braun JM, Meeker JD, Cooper GS. Phthalate exposure and male reproductive outcomes: A systematic review of the human epidemiological evidence. Environ Int. 2018 Dec;121(Pt 1):764-793.
[233] Yost EE, Euling SY, Weaver JA, Beverly BEJ, Keshava N, Mudipalli A, Arzuaga X, Blessinger T, Dishaw L, Hotchkiss A, Makris SL. Hazards of diisobutyl phthalate (DIBP) exposure: A systematic review of animal toxicology studies. Environ Int. 2019 Apr;125:579-594.
[234] U.S. EPA (2023). Draft Proposed Approach for Cumulative Risk Assessment of High-Priority Phthalates and a Manufacturer-Requested Phthalate under the Toxic Substances Control Act, p. 102.
[235] U.S. EPA (2020). Risk evaluation for cyclic aliphatic bromide cluster (HBCD).
[236] U.S. EPA (2021). Human health toxicity values for perfluorobutane sulfonic acid and related compound potassium perfluorobutane sulfonate. https://cfpub.epa.gov/ncea/risk/recordisplay.cfm?deid=350888.
[237] U.S. EPA (2020). Risk Evaluation for Percholorethylene (Ethene, 1,1,2,2-Tetrachloro-).
[238] U.S. EPA (2020). Risk Evaluation for Percholorethylene (Ethene, 1,1,2,2-Tetrachloro-).
[239] U.S. EPA (2020). Risk Evaluation for Trichloroethylene.
[240] U.S. EPA (2020). Risk Evaluation for Percholorethylene (Ethene, 1,1,2,2-Tetrachloro-).
[241] Radke EG, Braun JM, Meeker JD, Cooper GS. Phthalate exposure and male reproductive outcomes: A systematic review of the human epidemiological evidence. Environ Int. 2018 Dec;121(Pt 1):764-793.
[242] Yost EE, Euling SY, Weaver JA, Beverly BEJ, Keshava N, Mudipalli A, Arzuaga X, Blessinger T, Dishaw L, Hotchkiss A, Makris SL. Hazards of diisobutyl phthalate (DIBP) exposure: A systematic review of animal toxicology studies. Environ Int. 2019 Apr;125:579-594.
[243] U.S. EPA (2006). Organophosphorus cumulative risk assessment. https://www.regulations.gov/document/EPA-HQ-OPP-2006-0618-0002.
[244] U.S. EPA (2008). Revised N-methyl carbamate cumulative risk assessment. https://www.regulations.gov/document/EPA-HQ-OPP-2008-0347-0029.

EPA must either document that it has considered outcomes like altered thyroid hormone levels and other biochemical changes or cellular-level effects to be included in the animal and human evidence streams in the DINP Draft Hazard Assessment, or provide a justification for why these outcomes should not be considered as potential hazards of DINP.

Tagging biochemical and cellular-level outcomes as "supplemental, mechanistic," as directed in the PECO statement above, constrains the role of biochemical outcomes and other cellular changes to possibly providing biological support for apical outcomes, rather than considering precursors to apical outcomes as critical effects. Further, under EPA's proposed method, if no studies have been conducted of apical outcomes related to a biochemical outcome that has been studied, it is unclear whether the biochemical outcome will be considered at all. EPA says that supplemental studies "**may** be reviewed, evaluated for data quality, and incorporated into risk evaluations **as needed** for each chemical assessment"[245] (emphasis added), but it is unclear how a determination would be made to incorporate these studies into the risk evaluation, particularly in the absence of a related apical outcome study. Even if included to support a hazard conclusion based on apical outcomes, it appears that EPA rules out considering such studies for deriving a point of departure.

Exclusive reliance on studies of apical endpoints is also inconsistent with the best available science.[246] An important theme of the NASEM 2007 *Toxicity Testing in the 21st Century* report was that toxicity testing should move away from reliance on testing of apical outcomes. Accordingly, EPA's research programs and other U.S. health agencies have invested heavily in this new direction. Government and academic toxicology labs now rarely conduct studies of apical endpoints because the science has shifted towards examining more sensitive endpoints representing upstream biological changes ("key events") that lead to apical outcomes. In addition, a restriction to consider only apical or organ-level studies may bias the evidence base of the TSCA risk evaluations toward inclusion of industry-funded guidelines studies that are generally focused on apical endpoints.

e. EPA's approach to evidence integration lacks clear procedures and clearly-stated conclusions regarding the hazards of DINP.

EPA's TSCA risk evaluations lack a transparent and consistent approach to evidence integration. A key objective of the evidence integration process is to succinctly summarize the strength of the evidence concerning specific health endpoints and outcomes. This objective is advanced by pre-specifying a standard set of evidence descriptors. EPA's IRIS program handbook outlines a clear and consistent set of procedures for evidence synthesis and evidence integration that are applied in all IRIS assessments. The IRIS structure process culminates in selection of a concise descriptor summarizing the strength of evidence for each hazard – selected from the standardized terms "evidence demonstrates," "evidence indicates," and "evidence suggests" as hazard conclusions.[247] No such terms are used in DINP draft non-cancer hazard assessment DINP Draft

[245] U.S. EPA (2021). Draft systematic review protocol supporting TSCA risk evaluations for chemical substances, p. 345.
[246] 15 U.S.C. § 2625(h).
[247] U.S. EPA (2022). ORD Staff Handbook for Developing IRIS Assessments, Table 6-7.

Hazard Assessment. The TSCA risk evaluations do not demonstrate a consistent structured process to evidence integration, and concise phrases to summarize the evidence are not standardized and vary significantly within and across risk evaluations. The hazard conclusions for DINP uses concise but inconsistent and undefined phrases for some hazards, and only longer ambiguous phrases for other hazards. Summary terms used by EPA for DINP hazards include:

- "consistent evidence"[248] (liver toxicity)
- "Some evidence"[249] (neurotoxicity)
- "limited evidence"[250] (cardiovascular health effects and musculoskeletal toxicity)
- "DINP has consistently been shown to cause developmental effects in animal models"[251]
- "DINP lacks estrogenic potential *in vivo*"[252]
- For kidney toxicity, no phrasing representing an overall conclusion is provided; the clearest attempt at summarizing the evidence says "Findings were similar across study designs… EPA is considering kidney toxicity for dose-response analysis."[253]
- For immune system toxicity, no phrasing representing an overall conclusion is provided; the clearest attempt at summarizing the evidence says "Although available studies of laboratory animals provide evidence for immune adjuvant effects of DINP in sensitized animals, EPA is not further considering these effects for dose-response assessment or for use in extrapolating human risk."[254]

Without further standardization and definition of terms, it is difficult for readers to gain a clear, concise understanding of EPA's hazard conclusions. It is unclear, for example, if "consistent evidence" is equivalent to "strong evidence," or whether "findings were similar across study designs" is equivalent to "consistent evidence."

EPA should adopt a standardized procedure, such as the approach used by the IRIS program, for evidence integration for all DINP endpoints, including a pre-specified set of descriptors that are considered for each endpoint.

3. EPA's assessment of DINP carcinogenicity failed to recognize mechanisms in addition to PPARα activation that can contribute to animal liver tumors.

In assessing the carcinogenicity of DINP, EPA relies on an interpretation that all observed rodent tumors are attributable to the peroxisome proliferator–activated receptor-α (PPARα) pathway. EPA says:

[248] U.S. EPA (2024). Draft Non-cancer Human Health Hazard Assessment for Diisononyl Phthalate (DINP), p. 43.
[249] U.S. EPA (2024). Draft Non-cancer Human Health Hazard Assessment for Diisononyl Phthalate (DINP), p. 58.
[250] U.S. EPA (2024). Draft Non-cancer Human Health Hazard Assessment for Diisononyl Phthalate (DINP), p. 61 and p. 68.
[251] U.S. EPA (2024). Draft Non-cancer Human Health Hazard Assessment for Diisononyl Phthalate (DINP), p. 42.
[252] U.S. EPA (2024). Draft Non-cancer Human Health Hazard Assessment for Diisononyl Phthalate (DINP), p. 42.
[253] U.S. EPA (2024). Draft Non-cancer Human Health Hazard Assessment for Diisononyl Phthalate (DINP), pp. 50-51.
[254] U.S. EPA (2024). Draft Non-cancer Human Health Hazard Assessment for Diisononyl Phthalate (DINP), p. 67.

> Under the *Guidelines for Carcinogen Risk Assessment* (U.S. EPA, 2005), EPA reviewed the weight of evidence and determined that DINP is *Not Likely to be Carcinogenic to Humans* at doses below levels that do not result in PPARα activation…the non-cancer chronic POD of 15 mg/kg-day is considered protective of PPARα activation.[255]

However, experimental evidence for di(2-ethylhexyl)phthalate (DEHP) indicates that liver tumors in phthalate toxicity studies can occur by mechanisms other than PPARα activation:

> The recent study by Ito et al. (2007a) suggests that DEHP can induce PPAR-α–independent tumors without any loss of potency. In addition, as demonstrated by Yang et al. (2007) in their transgenic model of PPAR-α activation in hepatocytes, a robust hepatocyte and peroxisome proliferative response is itself insufficient to cause tumorigenesis. Despite their potential limitations, **these studies cast doubt on whether the proposed "key events" such as hepatocyte proliferation play a causal role in tumorigenesis or are merely correlated with cancer…A recent review (Rusyn et al. 2006) addressed other mechanistic effects of DEHP and proposed that tumors arise from a combination of molecular signals and pathways, rather than from a single event such as PPAR-α activation**.[256] (emphasis added)

EPA does not address many points raised by the California Office of Environmental Health Hazard Assessment (OEHHA) regarding evidence that liver tumors in DINP-exposed rodents may not be attributable to PPARα activation, including:

> **The degree to which a PPARα agonist induces liver responses indicative of PPARα activation has not been found to correlate with the ability of the chemical to induce liver tumors.** For example, in the F344 rat studies of Smith *et al.* (2000) described above, DIDP was found to be a more potent inducer of hepatic PCoA activity [a measure of PPARα activation] than DINP (by approximately 3-fold), yet no liver tumors were observed in studies of male and female F344 rats exposed to DIDP in the diet for two years (0, 400, 2000 or 8000 parts per million).[257] (emphasis added)
>
> The inconsistency of the short-term hepatocellular proliferation in DINP-exposed rats and mice and the lack of sustained long-term hepatocellular proliferation in DINP-exposed rats suggests that **PPARα activation may not be causally related to DINP-induced liver tumors in rats and mice**.[258] (emphasis added)
>
> DINP induces a number of liver changes in rodents consistent with PPARα activation. However, **studies with PPARα-null mice indicate that DINP induces some of these liver changes independently of PPARα activation** (e.g., increased LBWR associated with older age at exposure in female PPARα-null mice and increased β/ω fatty acid

[255] U.S. EPA (2024). Draft Cancer Human Health Hazard Assessment for Diisononyl Phthalate (DINP), p. 38.
[256] Guyton KZ, Chiu WA, Bateson TF, Jinot J, Scott CS, Brown RC, Caldwell JC. A reexamination of the PPAR-alpha activation mode of action as a basis for assessing human cancer risks of environmental contaminants. Environmental Health Perspectives, Nov 2009, (117)11:1664 – 1672.
[257] OEHHA (2013). Evidence of the carcinogenicity of diisononyl phthalate (DINP), p. 54. https://oehha.ca.gov/media/downloads/proposition-65/chemicals/dinphid100413.pdf.
[258] OEHHA (2013). Evidence of the carcinogenicity of diisononyl phthalate (DINP), p. 56.

> oxidation enzyme induction associated with younger age at exposure in male and female PPARα-null mice).[259] (emphasis added)

In addition, the Internation Agency for Research on Cancer (IARC) 2011 assessment of DEHP similarly concludes that rodent liver tumors were not necessarily due to the PPARα pathway.[260] In a previous review completed in 2000, IARC had concluded that DEHP was "not classifiable" for cancer in humans based on an interpretation that the liver tumors observed in rodent studies were attributable to PPARα activation and were judged not relevant to humans.[261] In 2011, IARC updated its assessment and came to a different conclusion. IARC reported that newer studies had been conducted in multiple varieties of genetically-engineered mice:

> These provide important additional data key for consideration of the relevance of the PPARα mode of action to rodent and human liver carcinogenesis. These include, but are not limited to, studies in *Pparα*-null mice, *PPARα* humanized transgenic mice and hepatocyte-specific constitutively activated *Pparα* transgenic mice (Yang *et al.*, 2007). The data from these animal models suggest that, although the activation of PPARα and the subsequent downstream events mediated by this transcription factor represent one key mechanism of action, **it is evident that several additional molecular signals and multiple pathways in several cell types in the liver, rather than a single molecular event, contribute to the formation of liver tumours in rats and mice.[262]** (emphasis added)
>
> Since the previous evaluation, important additional mechanistic information has become available, including, but not limited to, subacute, subchronic and chronic studies with di(2-ethylhexyl) phthalate in peroxisome proliferator-activated receptor α-null mice, as well as findings from several transgenic (peroxisome proliferator-activated receptor α-humanized and hepatocyte-specific constitutively activated peroxisome proliferator-activated receptor α mouse lines. Activation of peroxisome proliferator-activated receptor α and the subsequent downstream events mediated by this transcription factor represent an important mechanism of action for di(2-ethylhexyl) phthalate in rats and mice. However, additional data from animal models and studies in humans exposed to di(2-ethylhexyl) phthalate from the environment suggest that multiple molecular signals and pathways in several cell types in the liver, rather than a single molecular event, contribute to the induction of cancer in rats and mice. Thus, the relevance to human cancer of the molecular events that lead to cancer elicited by di(2-ethylhexyl) phthalate in several target tissues (e.g. the liver and testis) in rats and mice cannot be ruled out.[263]

These findings for DEHP provide important information that must inform EPA's interpretation regarding the PPARα pathway for phthalate carcinogenicity, and must be incorporated into the

---

[259] OEHHA (2013). Evidence of the carcinogenicity of diisononyl phthalate (DINP), pp. 65-66.
[260] IARC (2011). Some Chemicals Present in Industrial and Consumer Products, Food and Drinking-water. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans Volume 101, p. 254.
[261] IARC (2000). Di(2-ethylhexyl) phthalate (DEHP). IARC Monographs Vol 77.
[262] IARC (2011). Some Chemicals Present in Industrial and Consumer Products, Food and Drinking-water. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans Volume 101, p. 254.
[263] IARC (2011). Some Chemicals Present in Industrial and Consumer Products, Food and Drinking-water. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans Volume 101, p.259.

DINP Draft Hazard Assessment. IARC's 2011 update concluded that DEHP is "possibly carcinogenic to humans," which corresponds to an EPA cancer guidelines descriptor of "Suggestive Evidence of Carcinogenic Potential." The carcinogenicity evidence base for DINP is similar to DEHP, so an appropriate descriptor for DINP is at least "suggestive evidence."

Further, EPA's assertion that DINP liver tumors have a threshold dose-response relationship, with common levels of exposure posing no cancer risk, is inconsistent with the best available science. Given all the evidence reviewed above, EPA cannot conclude that liver tumors occur only through a threshold mechanism. Further, in *Science and Decisions,* the National Academies recommended that a conclusion of a threshold dose-response relationship cannot be made without first considering broad differences in human variability, background exposure, and background biology:

> The current EPA practice of determining "nonlinear" MOAs does not account for mechanistic factors that can create linearity at low dose. The dose-response relationship can be linear at a low dose when an exposure contributes to an existing disease process (Crump et al. 1976, Lutz 1990). Effects of exposures that add to background processes and background endogenous and exogenous exposures can lack a threshold if a baseline level of dysfunction occurs without the toxicant and the toxicant adds to or augments the background process. Thus, even small doses may have a relevant biologic effect. That may be difficult to measure because of background noise in the system but may be addressed through dose-response modeling procedures. Human variability with respect to the individual thresholds for a nongenotoxic cancer mechanism can result in linear dose-response relationships in the population (Lutz 2001).
> In the laboratory, nonlinear dose-response processes—for example, cytotoxicity, impaired immune function and tumor surveillance, DNA methylation, endocrine disruption, and modulation of cell cycles—may be found to cause cancer in test animals. However, given the high prevalence of those background processes, given cancer as an end point, and given the multitude of chemical exposures and high variability in human susceptibility, the results may still be manifested as low-dose linear dose-response relationships in the human population (Lutz 2001).[264]

Science and Decisions further stated that, before concluding that a threshold dose-response relationship is operable, EPA should undertake:

> Systematic assessments of the MOAs, vulnerable populations, and background exposures and disease processes that may affect a chemical's human dose-response relationships and human vulnerability. This includes an evaluation of the potential background exposures and processes (for example, damage and repair processes, disease, and aging) that interact with a chemical's MOAs and thus contribute to variability in and vulnerability to the toxicant response and that can result in a population dose-response relationship that is linear at low doses.[265]

[264] National Research Council (2009). Science and Decisions: Advancing Risk Assessment, pp. 129-130.
[265] National Research Council (2009). Science and Decisions: Advancing Risk Assessment, p. 178.

EPA has not considered the issues raised by *Science and Decisions*, it has not conducted the recommended systematic assessments regarding the cancer dose-response relationship for DINP, and it has not considered the multiple lines of experimental evidence summarized above pointing to multiple mechanisms for DINP liver tumors. EPA's conclusion of a threshold is therefore not consistent with the best available science, and the addition of any language to the cancer descriptor concerning the dose at which PPARα activation occurs is scientifically inappropriate.

Technical Appendix A: Application of IPCS framework to DIDP non-cancer risks

In the *Draft Risk Evaluation for Diisodecyl Phthalate (DIDP)*, EPA selected developmental toxicity (reduced F2 offspring survival) for estimation of risks from chronic oral exposures. EPA correctly identifies this outcome as "clearly adverse."[1,2]

For risk characterization of non-cancer health effects, the draft TSCA risk evaluation calculates a "margin of exposure" (MOE) for each exposure scenario, which is the ratio of the point of departure (POD) to the exposure level. For the DIDP developmental effects, the *DIDP Draft Risk Evaluation* concludes that an MOE of 30 or more indicates that "risk is not considered to be of concern and mitigation is not needed."[3] EPA's approach to risk characterization does not actually estimate risks of adverse effects in the population with chronic exposure to DIDP, but instead simply applies a "bright line" judgment of whether or not the MOE is adequate. A more informative approach for both risk characterization and risk management would be to apply the probabilistic dose-response assessment methods of the International Programme on Chemical Safety (IPCS),[4] part of the World Health Organization (WHO), to estimate the risk of adverse effects at various levels of exposure. The IPCS methodology has previously been described and applied in several peer-reviewed journal articles.[5,6,7,8,9]

We applied the IPCS approach for "quantal-deterministic" endpoints and the "approximate probabilistic" calculation (see IPCS report Fig 3.5, panel C)[10] to estimate risks of reduced numbers of developmental toxicity from chronic oral exposure to DIDP. The analysis involved the following steps:

1. Derivation of IPCS POD and corresponding uncertainty adjustments
2. Application of interspecies adjustments
3. Application of intraspecies adjustments
4. Calculation of HD_M^I - the human dose (HD) of DIDP associated with a particular magnitude of effect M at a particular population incidence I.

For each aspect of the analysis, including the values used to derive the IPCS POD and the adjustment factors applied to derive the HD_M^I, the IPCS methodology uses a 50th percentile value (P50) as a central estimate and the ratio of 95th percentile to 50th percentile (P95/P50) as a measure of uncertainty. All POD and HD_M^I values presented in this analysis are for continuous exposures.

We demonstrate each of these steps starting with the EPA-determined no-observed-adverse-effect level (NOAEL) (i.e. the chronic oral POD in applied dose units) to derive a set of oral HD_M^I values for different levels of population incidence. Use of a benchmark dose (BMD) is preferred as the starting point for application of the WHO/IPCS framework, but EPA did not conduct BMD modeling for the developmental toxicity endpoint or other candidate PODs.

STEP 1: Derivation of IPCS POD and corresponding uncertainty adjustments

The IPCS methodology requires the use of an ED_{50} (median effective dose) value as the POD for quantal-deterministic endpoints. Since an ED_{50} is not available from the EPA risk evaluation, we began with EPA's NOAEL, which 38 mg/kg-day, and applied adjustments provided by the IPCS methodology. Uncertainty in the NOAEL estimate is unquantified, which is represented by a P95/P50 value of 1.

To estimate an ED_{50} from a NOAEL for a quantal-deterministic developmental toxicity endpoint, the IPCS framework multiplies the NOAEL by 2/9 (central estimate, or P50) with uncertainty (P95/P50) equal to 5.0.[11] The adjustment from the NOAEL to ED_{50} to derive the IPCS POD is entered in the IPCS approximate probabilistic calculation template as follows:

Determination of point of departure (POD) and its uncertainty[a] for probabilistic dose-response analysis of chronic oral exposure to DIDP: reduced offspring survival		
Aspect	**P50**	**P95/P50**
NOAEL	38 mg/kg-d	1[b]
NOAEL-to-ED_{50} adjustment	0.22[c]	5[c]
IPCS POD = ED_{50}	**171 mg/kg-d[d]**	**5[e]**

[a] Uncertainty is expressed as the ratio of the 95th percentile (P95) to the 50th percentile (P50)
[b] No estimate of uncertainty for the NOAEL is available, so P/95/P50 = 1
[c] IPCS Table 4.1: quantal (deterministic) - developmental effects
[d] IPCS POD = (38 mg/kg-d) / 0.22 = 171 mg/kg-d
[e] (Composite P95/P50) = $10^{[(\log 1)^2 + (\log 5)^2]^{0.5}} = 5$

Step 2: Application of interspecies (animal-to-human) adjustments

For interspecies (animal-to-human) adjustments, the IPCS methodology first considers a factor for body-size scaling, and then a factor for remaining toxicokinetic (TK) and toxicodynamic (TD) differences. The IPCS framework provides equations for calculating the body size scaling adjustment factor, based on the assumed human body weight and test species body weight. We applied EPA's assumptions for human body weight (80 kg) and rat body weight (0.25 kg)[12] to derive the appropriate central tendency (P50) factor and its uncertainty (P95/P50).

For the TK/TD differences remaining after bodyweight scaling, the IPCS report recommends a central estimate (P50) of 1 (i.e., no additional interspecies differences) and representing uncertainty with a P95/P50 factor of 3.[13] We incorporated these IPCS recommendations, which are entered In the IPCS approximate probabilistic calculation template as follows:

Interspecies adjustments for probabilistic dose-response analysis of Chronic oral exposure to DIDP: reduced offspring survival		
Aspect	**P50**	**P95/P50**
$AF_{Interspecies\text{-}BS}$	5.64[a]	1.26[a]
$AF_{Interspecies\text{-}TK/TD}$	1	3

[a] Calculated from IPCS equation 4-2 using EPA assumptions regarding body weight of humans (80 kg) and rats (0.25 kg).

Step 3: Application of intraspecies (human variability) adjustments

In the IPCS methodology, the value of the human variability adjustment factor ($AF_{intraspecies}$) varies depending on the incidence of the adverse effect in the exposed population – with a larger adjustment factor necessary to extrapolate from the POD to lower levels of incidence. The IPCS report provides $AF_{intraspecies}$ for several incidence (I) values. The P50 and P95/P50 values for $AF_{intraspecies}$ provided by IPCS for several values of I, along with additional values of I of interest for this analysis, are provided in the following table:

Lognormal approximation of uncertainty distributions for intraspecies variability ($AF_{Intraspecies}$) for varying levels of population incidence (I)		
Incidence (I)	**$AF_{Intraspecies}$**	
	P50	**P95/P50**
10%[a]	3.49	2.24
5%[a]	4.98	2.82
1%[a]	9.69	4.32
0.1% (1-in-1,000)[a]	20.42	6.99
0.01% (1-in-10,000)[a]	37.71	10.39
0.001% (1-in-100,000)[b]	64.25	14.65

[a] IPCS Table 4.5
[b] Calculated for this analysis using the same methods that were used to derive IPCS Table 4.5

Step 4: Calculation of HD_M^I

The output of the IPCS methodology is generically described as an HD_M^I value – the human dose (HD) associated with a particular magnitude of effect M at a particular population incidence I. For this analysis, the "M" represents the developmental effect of reduced offspring survival. The following tables present the HD_M^I results for I = 10%, 5%, 1%, 0.1%, 0.01%, and 0.001% using the POD, $AF_{Interspecies}$, and $AF_{Intraspecies}$ values shown above.
The IPCS approach is a probabilistic method, so the HD_M^I is a distribution; selected values from that distribution are presented in the tables as follows:

- P05: 5th percentile estimate (lower confidence limit) of HD_M^I (this value is shown in **bold**)
- P50: 50th percentile estimate (median) of HD_M^I
- P95: 95th percentile estimate (upper confidence limit) of HD_M^I.

All HD_M^I values in the following tables are human equivalent doses (HEDs), as they incorporate interspecies body size scaling (see Step 2 above).

Calculation of HD_M^I for chronic oral exposure to DIDP: reduced offspring survival (Incidence = 10%)		
Aspect	**P50**	**P95/P50**
NOAEL	38 mg/kg-d	1
NOAEL-to-ED_{50} adjustment	0.22	5
IPCS POD = ED_{50}	**171 mg/kg-d**	**5**
$AF_{Interspecies-BS}$	5.64	1.26
$AF_{Interspecies-TK/TD}$	1	3
$AF_{Intra-I=10\%}$	3.49	2.24
HD_M^I	8.7 mg/kg-d[a]	8.3[b]
	P05	**P95**
HD_M^I [(c)]	**1.0 mg/kg-d**	72 mg/kg-d

[a] HD_M^I (P50) = IPCS POD / ($AF_{Interspecies-BS}$ x $AF_{Interspecies-TK/TD}$ x $AF_{Intraspecies}$)
[b] (Composite P95/P50) = $10^{[(\log 5)^2 + (\log 1.26)^2 + (\log 3)^2 + (\log 2.24)^2]^{0.5}} = 8.3$
[c] HD_M^I (P05) = HD_M^I (P50) / (Composite P95/P50)
HD_M^I (P95) = HD_M^I (P50) x (Composite P95/P50)

Calculation of HD_M^I for chronic oral exposure to DIDP: reduced offspring survival (Incidence = 5%)		
Aspect	**P50**	**P95/P50**
NOAEL	38 mg/kg-d	1
NOAEL-to-ED_{50} adjustment	0.22	5
IPCS POD = ED_{50}	**171 mg/kg-d**	**5**
$AF_{Interspecies-BS}$	5.64	1.26
$AF_{Interspecies-TK/TD}$	1	3
$AF_{Intra-I=5\%}$	4.98	2.82
HD_M^I	6.1 mg/kg-d[a]	9.2[b]
	P05	**P95**
$HD_M^{I\,(c)}$	**0.66 mg/kg-d**	56 mg/kg-d

[a] HD_M^I (P50) = IPCS POD / ($AF_{Interspecies-BS}$ x $AF_{Interspecies-TK/TD}$ x $AF_{Intraspecies}$)
[b] (Composite P95/P50) = $10^{[(\log 5)^2 + (\log 1.26)^2 + (\log 3)^2 + (\log 2.82)^2]^{0.5}} = 9.2$
[c] HD_M^I (P05) = HD_M^I (P50) / (Composite P95/P50)
HD_M^I (P95) = HD_M^I (P50) x (Composite P95/P50)

Calculation of HD_M^I for chronic oral exposure to DIDP: reduced offspring survival (Incidence = 1%)		
Aspect	**P50**	**P95/P50**
NOAEL	38 mg/kg-d	1
NOAEL-to-ED_{50} adjustment	0.22	5
IPCS POD = ED_{50}	**171 mg/kg-d**	**5**
$AF_{Interspecies-BS}$	5.64	1.26
$AF_{Interspecies-TK/TD}$	1	3
$AF_{Intra-I=1\%}$	9.69	4.32
HD_M^I	3.1 mg/kg-d[a]	11.6[b]
	P05	**P95**

HD_M^{I} [(c)]	**0.27 mg/kg-d**	36 mg/kg-d

[a] HD_M^{I} (P50) = IPCS POD / ($AF_{Interspecies-BS}$ x $AF_{Interspecies-TK/TD}$ x $AF_{Intraspecies}$)
[b] (Composite P95/P50) = $10^{[(\log 5)^2 + (\log 1.26)^2 + (\log 3)^2 + (\log 4.32)^2]^{0.5}} = 11.6$
[c] HD_M^{I} (P05) = HD_M^{I} (P50) / (Composite P95/P50)
HD_M^{I} (P95) = HD_M^{I} (P50) x (Composite P95/P50)

Calculation of HD_M^{I} for chronic oral exposure to DIDP: reduced offspring survival (Incidence = 0.1%)

Aspect	P50	P95/P50
NOAEL	38 mg/kg-d	1
NOAEL-to-ED_{50} adjustment	0.22	5
IPCS POD = ED_{50}	**171 mg/kg-d**	**5**
$AF_{Interspecies-BS}$	5.64	1.26
$AF_{Interspecies-TK/TD}$	1	3
$AF_{Intra-I=0.1\%}$	20.42	6.99
HD_M^{I}	1.5 mg/kg-d[a]	15.8[b]
	P05	**P95**
HD_M^{I} [(c)]	**0.1 mg/kg-d**	24 mg/kg-d

[a] HD_M^{I} (P50) = IPCS POD / ($AF_{Interspecies-BS}$ x $AF_{Interspecies-TK/TD}$ x $AF_{Intraspecies}$)
[b] (Composite P95/P50) = $10^{[(\log 5)^2 + (\log 1.26)^2 + (\log 3)^2 + (\log 6.99)^2]^{0.5}} = 15.8$
[c] HD_M^{I} (P05) = HD_M^{I} (P50) / (Composite P95/P50)
HD_M^{I} (P95) = HD_M^{I} (P50) x (Composite P95/P50)

Calculation of HD_M^{I} for chronic oral exposure to DIDP: reduced offspring survival (Incidence = 0.01%)

Aspect	P50	P95/P50
NOAEL	38 mg/kg-d	1
NOAEL-to-ED_{50} adjustment	0.22	5
IPCS POD = ED_{50}	**171 mg/kg-d**	**5**
$AF_{Interspecies-BS}$	5.64	1.26

$AF_{Interspecies\text{-}TK/TD}$	1	3
$AF_{Intra\text{-}I=0.01\%}$	37.71	10.39
HD_M^I	0.8 mg/kg-d[a]	21.2[b]
	P05	**P95**
HD_M^I [(c)]	**0.04 mg/kg-d**	17 mg/kg-d

[a] HD_M^I (P50) = IPCS POD / ($AF_{Interspecies\text{-}BS}$ x $AF_{Interspecies\text{-}TK/TD}$ x $AF_{Intraspecies}$)
[b] (Composite P95/P50) = $10^{[(\log 5)^2 + (\log 1.26)^2 + (\log 3)^2 + (\log 10.39)^2]^{0.5}} = 21.2$
[c] HD_M^I (P05) = HD_M^I (P50) / (Composite P95/P50)
HD_M^I (P95) = HD_M^I (P50) x (Composite P95/P50)

Calculation of HD_M^I for chronic oral exposure to DIDP: reduced offspring survival (Incidence = 0.001%)		
Aspect	**P50**	**P95/P50**
NOAEL	38 mg/kg-d	1
NOAEL-to-ED_{50} adjustment	0.22	5
IPCS POD = ED_{50}	**171 mg/kg-d**	**5**
$AF_{Interspecies\text{-}BS}$	5.64	1.26
$AF_{Interspecies\text{-}TK/TD}$	1	3
$AF_{Intra\text{-}I=0.001\%}$	64.25	14.65
HD_M^I	0.5 mg/kg-d[a]	27.8[b]
	P05	**P95**
HD_M^I [(c)]	**0.02 mg/kg-d**	13 mg/kg-d

[a] HD_M^I (P50) = IPCS POD / ($AF_{Interspecies\text{-}BS}$ x $AF_{Interspecies\text{-}TK/TD}$ x $AF_{Intraspecies}$)
[b] (Composite P95/P50) = $10^{[(\log 5)^2 + (\log 1.26)^2 + (\log 3)^2 + (\log 14.65)^2]^{0.5}} = 27.8$
[c] HD_M^I (P05) = HD_M^I (P50) / (Composite P95/P50)
HD_M^I (P95) = HD_M^I (P50) x (Composite P95/P50)

Interpretation of Results

The National Academies and the WHO/IPCS have both recommended using the lower confidence limit (LCL) on a probabilistic dose-response distribution for use in decision-making,

in place of a traditional reference dose (RfD) or reference concentration (RfC). The National Academies said in *Science and Decisions* that:

> Multiple risk-specific doses could be provided…in the various risk characterizations that EPA produces to aid environmental decision-making.[14]
>
> A Risk-Specific Reference Dose: For quantal effects, the RfD can be defined to be the dose that corresponds to a particular risk specified to be de minimis (for example, 1 in 100,000) at a defined confidence level (for example, 95%) for the toxicity end point of concern.[15]

The WHO/IPCS said:

> The LCL of the HD_M^I can be used as a probabilistic RfD to replace the deterministic RfD. In this case, the probabilistic RfD is the dose that protects the population from a specified magnitude and incidence of effect with a pre-specified per cent coverage (confidence).[16]

Consistent with the guidance from the National Academies and the IPCS, we summarize the above results by focusing on the lower confidence limit (5th percentile or P05) risk-specific doses (HD_M^I) for multiple levels of risk (incidence or I).

Based on application of the WHO/IPCS methodology to DIDP developmental effects from chronic exposures, we find that:

1.0 mg/kg-day is the lower bound (95% confidence) chronic human dose at which clearly adverse developmental effects are expected in 10% of the exposed population

0.27 mg/kg-day is the lower bound (95% confidence) chronic human dose at which clearly adverse developmental effects are expected in 1% of the exposed population

0.1 mg/kg-day is the lower bound (95% confidence) chronic human dose at which clearly adverse developmental effects are expected in 0.1% of the exposed population

0.04 mg/kg-day is the lower bound (95% confidence) chronic human dose at which clearly adverse developmental effects are expected in 0.01% (1-in-10,000) of the exposed population

0.02 mg/kg-day is the lower bound (95% confidence) chronic human dose at which clearly adverse developmental effects are expected in 0.001% (1-in-100,000) of the exposed population

The implications of these risk values can be understood by comparison with the exposure levels considered by EPA to represent negligible risk. EPA's assessment uses a POD of 9.0 mg/kg-day and a benchmark MOE of 30, meaning that EPA concludes "risk is not considered to be of concern and mitigation is not needed"[17] for any exposure below 0.30 mg/kg-day (9.0 mg/kg-day / 30 = 0.30 mg/kg-day). Our analysis indicates that an exposure of 0.30 mg/kg-day exceeds the

lower-bound dose for the 1% (1-in-100) risk level. This risk far exceeds EPA's usual target range of protection for carcinogenic risks of 1-in-10,000 to 1-in-1,000,000.[18]

The risk values obtained from application of the WHO framework also indicate that many workers are at high risk for adverse non-cancer effects:

- High-end exposure estimates for 3 occupational exposure scenarios are greater than or equal to 0.27 mg/kg, the lower-bound dose estimate for 1% (1-in-100) risk: application of adhesives and sealants, application of paints and coatings, and use of penetrants and inspection fluids
- High-end exposure estimates for an additional 5 occupational exposure scenarios are greater than or equal to 0.1 mg/kg, the lower-bound dose estimate for 0.1% (1-in-1,000) risk: PVC plastics compounding, PVC plastics converting, non-PVC material converting, recycling, and disposal
- Central-tendency exposure estimates for 5 occupational exposure scenarios are greater than or equal to 0.04 mg/kg, the lower-bound dose estimate for 0.01% (1-in-10,000) risk: application of adhesives and sealants, application of paints and coatings, use of penetrants and inspection fluids, PVC plastics compounding, and non-PVC material converting.

The estimates of HD_M^I presented here were based entirely on input values and equations available from the WHO/IPCS methodology document and related publications, and from EPA's *Draft Human Health Hazard Assessment for Diisodecyl Phthalate (DIDP)*. An important caveat to these calculations is that the values used to represent human variability may be understated. The IPCS default human variability distribution is based on 37 data sets for human toxicokinetic variability and 34 data sets for human toxicodynamic variability. Most of these data sets were obtained from controlled human exposure studies of pharmaceuticals conducted in small samples of healthy adults, representing considerably less variability than found in the general population.[19,20,21] If human variability is underestimated, then the actual dose associated with each incidence level (e.g. I =1%, I = 0.1%) will be lower than the values obtained from this analysis – or in other words, risk at each dose will be underestimated.

Technical Appendix B: Application of IPCS framework to DINP non-cancer risks

In the *Draft Non-cancer Human Health Hazard Assessment for Diisononyl Phthalate (DINP)*, EPA selected liver toxicity as the most sensitive endpoint for estimation of risks from chronic oral exposures.

For risk characterization of non-cancer health effects, TSCA risk evaluations calculate a "margin of exposure" (MOE) for each exposure scenario, which is the ratio of the point of departure (POD) to the exposure level. For the DINP liver effects, the DINP draft hazard assessment concludes that a benchmark MOE of 30[1] indicates that "risk is not considered to be of concern and mitigation is not needed."[2] EPA's approach to risk characterization does not actually estimate risks of adverse effects in the population, but instead simply applies a "bright line" judgment of whether or not the MOE is adequate. A more informative approach for both risk characterization and risk management would be to apply the probabilistic dose-response assessment methods of the International Programme on Chemical Safety (IPCS),[3] part of the World Health Organization (WHO), to estimate the risk of adverse effects at various levels of exposure. The IPCS methodology has previously been described and applied in several peer-reviewed journal articles.[4,5,6,7,8]

We applied the IPCS approach for "quantal-deterministic" and continuous endpoints and the "approximate probabilistic" calculation (see IPCS report Fig 3.5, panel C)[9] to estimate risks of two DINP liver toxicity endpoints: spongiosis hepatis (a type of liver lesion) and increased serum ALT (a biomarker indicating liver damage).

The analysis involved the following steps:

1. Derivation of IPCS POD and corresponding uncertainty adjustments
2. Application of interspecies adjustments
3. Application of intraspecies adjustments
4. Calculation of HD_M^I - the human dose (HD) of DINP associated with a particular magnitude of effect M at a particular population incidence I.

For each aspect of the analysis, including the values used to derive the IPCS POD and the adjustment factors applied to derive the HD_M^I, the IPCS methodology uses a 50th percentile value (P50) as a central estimate and the ratio of 95th percentile to 50th percentile (P95/P50) as a measure of uncertainty. All POD and HD_M^I values presented in this analysis are for continuous exposures.

We demonstrate each of these steps starting with the EPA-estimated benchmark dose (BMD) values in applied dose units to derive a set of oral HD_M^I values for different levels of population incidence (e.g. 1%, 0.1%, etc.). Although EPA has selected a NOAEL for liver toxicity as the chronic POD for DINP rather than the statistically-estimated BMD, EPA guidance states that BMDs are preferable to NOAELs for characterizing dose-response relationships (see main comments above).

STEP 1: Derivation of IPCS POD and corresponding uncertainty adjustments

EPA conducted BMD modeling for several liver endpoints from a study by Lington et al. The two most sensitive endpoints were spongiosis hepatis and increased serum ALT at 6 month sacrifice. BMD results were as follows:

EPA benchmark dose modeling results for two liver toxicity endpoints from LIngton et al. 1977			
Endpoint	**Benchmark Response (BMR)**	**BMD**	**BMDL**
Spongiosis hepatis[a]	10% relative deviation	31.88 mg/kg-d	8.57 mg/kg-d
Increased serum ALT at 6-month sacrifice[b]	1 standard deviation change from controls	23.42 mg/kg-d	15.50 mg/kg-d

[a] U.S. EPA (2024). Draft Non-cancer Human Health Hazard Assessment for Diisononyl Phthalate (DINP), Table_Apx E-15.
[b] U.S. EPA (2024). Draft Non-cancer Human Health Hazard Assessment for Diisononyl Phthalate (DINP), Table_Apx E-7.

In the IPCS methodology, the BMD is the central estimate (P50), and uncertainty in the BMD (P95/P50) is equal to the ratio of BMD / BMDL:

BMD/BMDL (spongiosis hepatis) = 31.88 / 8.57 = 3.72
BMD/BMDL (increased serum ALT) = 23.42 / 15.50 = 1.51.

In the IPCS methodology, spongiosis hepatis is classified as a quantal-deterministic endpoint. The IPCS methodology requires the use of an ED_{50} (median effective dose) value as the POD for quantal-deterministic endpoints. Since an ED_{50} is not available from the EPA risk evaluation, we began with the BMD, and applied adjustments provided by the IPCS methodology: "if ED50 not reported: BMD at the reported BMR is multiplied by an additional factor of 3.0; additional uncertainty through adding 1.5^2 to $(P95/P50)^2$."[10] For increased serum ALT, a continuous endpoint, no adjustment to the ED_{50} is applied and the BMD is used as the POD in applying the IPCS framework.

The values applied for determining the IPCS POD and its uncertainty for each endpoint are entered in the IPCS approximate probabilistic calculation template as follows:

Determination of point of departure (POD) and its uncertainty[a] for probabilistic dose-response analysis of chronic oral exposure to DINP: liver toxicity				
Aspect	**Spongiosis hepatis**		**Increased serum ALT**	
	P50	**P95/P50**	**P50**	**P95/P50**
BMD	31.9 mg/kg-d	3.72	23.4 mg/kg-d	1.51
BMD-to-ED_{50} adjustment	3	1.5	N/A[b]	N/A[b]
IPCS POD = ED_{50}	**95.6 mg/kg-d**	**3.95[c]**	**23.4 mg/kg-d**	**1.51**

[a] Uncertainty is expressed as the ratio of the 95th percentile (P95) to the 50th percentile (P50)
[b] Not applicable
[c] (Composite P95/P50) = $10^{[(\log 3.72)^2 + (\log 1.5)^2]^{0.5}} = 3.95$

Step 2: Application of interspecies (animal-to-human) adjustments

For interspecies (animal-to-human) adjustments, the IPCS methodology first considers a factor for body-size scaling, and then a factor for remaining toxicokinetic (TK) and toxicodynamic (TD) differences. The IPCS framework provides equations for calculating the body size scaling adjustment factor, based on the assumed human body weight and test species body weight. We applied EPA's assumptions for human body weight (80 kg) and rat body weight (0.25 kg)[11] to derive the appropriate central tendency (P50) factor and its uncertainty (P95/P50).

For the TK/TD differences remaining after bodyweight scaling, the IPCS report recommends a central estimate (P50) of 1 (i.e., no additional interspecies differences) and representing uncertainty with a P95/P50 factor of 3.[12] We incorporated these IPCS recommendations, which are entered In the IPCS approximate probabilistic calculation template as follows:

Interspecies adjustments for probabilistic dose-response analysis of chronic oral exposure to DINP: liver toxicity		
Aspect	**P50**	**P95/P50**
$AF_{Interspecies\text{-}BS}$	5.64[a]	1.26[a]
$AF_{Interspecies\text{-}TK/TD}$	1	3

[a] Calculated from IPCS equation 4-2 using EPA assumptions regarding body weight of humans (80 kg) and rats (0.25 kg).

Step 3: Application of intraspecies (human variability) adjustments

In the IPCS methodology, the value of the human variability adjustment factor ($AF_{intraspecies}$) varies depending on the incidence of the adverse effect in the exposed population – with a larger adjustment factor necessary to extrapolate from the POD to lower levels of incidence. The IPCS report provides $AF_{intraspecies}$ for several incidence (I) values. The P50 and P95/P50 values for $AF_{intraspecies}$ provided by IPCS for several values of I, along with additional values of I of interest for this analysis, are provided in the following table:

Lognormal approximation of uncertainty distributions for intraspecies variability ($AF_{Intraspecies}$) for varying levels of population incidence (I)		
Incidence (I)	**$AF_{Intraspecies}$**	
	P50	**P95/P50**
10%[a]	3.49	2.24
5%[a]	4.98	2.82
1%[a]	9.69	4.32
0.5% (1-in-200)[a]	12.36	5.06
0.1% (1-in-1,000)[a]	20.42	6.99
0.01% (1-in-10,000)[a]	37.71	10.39
0.001% (1-in-100,000)[b]	64.25	14.65

[a] IPCS Table 4.5
[b] Calculated for this analysis using the same methods that were used to derive IPCS Table 4.5

Step 4: Calculation of HD_M^I

The output of the IPCS methodology is generically described as an HD_M^I value – the human dose (HD) associated with a particular magnitude of effect M at a particular population incidence I. For this analysis, the "M" represents either spongiosis hepatis or increased serum ALT. The following tables present the HD_M^I results for I = 10%, 5%, 1%, 0.1%, 0.01%, and 0.001% using the POD, $AF_{Interspecies}$, and $AF_{Intraspecies}$ values shown above.

The IPCS approach is a probabilistic method, so the HD_M^I is a distribution; selected values from that distribution are presented in the tables as follows:

- P05: 5th percentile estimate (lower confidence limit) of HD_M^I (this value is shown in **bold**)
- P50: 50th percentile estimate (median) of HD_M^I
- P95: 95th percentile estimate (upper confidence limit) of HD_M^I.

All HD_M^I values in the following tables are human equivalent doses (HEDs), as they incorporate interspecies body size scaling (see Step 2 above).

Calculation of HD_M^I from chronic DINP exposure: liver toxicity (Incidence = 10%)				
Aspect	**Spongiosis hepatis**		**Increased serum ALT**	
	P50	**P95/P50**	**P50**	**P95/P50**
BMD	31.9 mg/kg-d	3.72	23.4 mg/kg-d	1.51
BMD-to-ED_{50} adjustment	3	1.5	N/A	N/A
IPCS POD = ED_{50}	**95.6 mg/kg-d**	**3.95[c]**	**23.4 mg/kg-d**	**1.51**
$AF_{Interspecies-BS}$	5.64	1.26	5.64	1.26
$AF_{Interspecies-TK/TD}$	1	3	1	3
$AF_{Intra-I=10\%}$	3.49	2.24	3.49	2.24
HD_M^I	4.9 mg/kg-d[a]	7.0[b]	1.2 mg/kg-d[a]	4.2[c]
	P05	**P95**	**P05**	**P95**
$HD_M^{I\,(d)}$	**0.7 mg/kg-d**	34 mg/kg-d	**0.3 mg/kg-d**	5 mg/kg-d

[a] HD_M^I (P50) = IPCS POD / ($AF_{Interspecies-BS}$ x $AF_{Interspecies-TK/TD}$ x $AF_{Intraspecies}$)
[b] (Composite P95/P50) = $10^{[(\log 3.95)^2 + (\log 1.26)^2 + (\log 3)^2 + (\log 2.24)^2]^{0.5}} = 7.0$
[c] (Composite P95/P50) = $10^{[(\log 1.51)^2 + (\log 1.26)^2 + (\log 3)^2 + (\log 2.24)^2]^{0.5}} = 4.2$
[d] HD_M^I (P05) = HD_M^I (P50) / (Composite P95/P50). HD_M^I (P95) = HD_M^I (P50) x (Composite P95/P50)

Calculation of HD_M^I from chronic DINP exposure: liver toxicity (Incidence = 5%)				
Aspect	**Spongiosis hepatis**		**Increased serum ALT**	
	P50	**P95/P50**	**P50**	**P95/P50**
BMD	31.9 mg/kg-d	3.72	23.4 mg/kg-d	1.51
BMD-to-ED_{50} adjustment	3	1.5	N/A	N/A
IPCS POD = ED_{50}	**95.6 mg/kg-d**	**3.95[c]**	**23.4 mg/kg-d**	**1.51**
$AF_{Interspecies-BS}$	5.64	1.26	5.64	1.26
$AF_{Interspecies-TK/TD}$	1	3	1	3
$AF_{Intra-I=5\%}$	4.98	2.82	4.98	2.82
HD_M^I	3.4 mg/kg-d[a]	7.8[b]	0.8 mg/kg-d[a]	4.9[c]
	P05	**P95**	**P05**	**P95**
$HD_M^{I\,(d)}$	**0.44 mg/kg-d**	27 mg/kg-d	**0.17 mg/kg-d**	4 mg/kg-d

[a] HD_M^I (P50) = IPCS POD / ($AF_{Interspecies-BS}$ x $AF_{Interspecies-TK/TD}$ x $AF_{Intraspecies}$)
[b] (Composite P95/P50) = $10^{[(\log 3.95)^2 + (\log 1.26)^2 + (\log 3)^2 + (\log 2.82)^2]^{0.5}} = 7.8$
[c] (Composite P95/P50) = $10^{[(\log 1.51)^2 + (\log 1.26)^2 + (\log 3)^2 + (\log 2.82)^2]^{0.5}} = 4.9$
[d] HD_M^I (P05) = HD_M^I (P50) / (Composite P95/P50). HD_M^I (P95) = HD_M^I (P50) x (Composite P95/P50)

Calculation of HD_M^I from chronic DINP exposure: liver toxicity (Incidence = 1%)				
Aspect	**Spongiosis hepatis**		**Increased serum ALT**	
	P50	**P95/P50**	**P50**	**P95/P50**
BMD	31.9 mg/kg-d	3.72	23.4 mg/kg-d	1.51
BMD-to-ED_{50} adjustment	3	1.5	N/A	N/A
IPCS POD = ED_{50}	**95.6 mg/kg-d**	**3.95[c]**	**23.4 mg/kg-d**	**1.51**
$AF_{Interspecies-BS}$	5.64	1.26	5.64	1.26
$AF_{Interspecies-TK/TD}$	1	3	1	3
$AF_{Intra-I=1\%}$	9.69	4.32	9.69	4.32
HD_M^I	1.7 mg/kg-d[a]	10.0[b]	0.04 mg/kg-d[a]	6.6[c]

	P05	P95	P05	P95
$HD_M^{I\,(d)}$	**0.18 mg/kg-d**	17 mg/kg-d	**0.06 mg/kg-d**	2.8 mg/kg-d

[a] HD_M^I (P50) = IPCS POD / ($AF_{Interspecies\text{-}BS}$ x $AF_{Interspecies\text{-}TK/TD}$ x $AF_{Intraspecies}$)
[b] (Composite P95/P50) = $10^{\wedge}[(\log 3.95)^2 + (\log 1.26)^2 + (\log 3)^2 + (\log 4.32)^2]^{0.5} = 10.0$
[c] (Composite P95/P50) = $10^{\wedge}[(\log 1.51)^2 + (\log 1.26)^2 + (\log 3)^2 + (\log 4.32)^2]^{0.5} = 6.6$
[d] HD_M^I (P05) = HD_M^I (P50) / (Composite P95/P50). HD_M^I (P95) = HD_M^I (P50) x (Composite P95/P50)

Calculation of HD_M^I from chronic DINP exposure: liver toxicity (Incidence = 0.5%)

Aspect	Spongiosis hepatis		Increased serum ALT	
	P50	P95/P50	P50	P95/P50
BMD	31.9 mg/kg-d	3.72	23.4 mg/kg-d	1.51
BMD-to-ED_{50} adjustment	3	1.5	N/A	N/A
IPCS POD = ED_{50}	**95.6 mg/kg-d**	**3.95[c]**	**23.4 mg/kg-d**	**1.51**
$AF_{Interspecies\text{-}BS}$	5.64	1.26	5.64	1.26
$AF_{Interspecies\text{-}TK/TD}$	1	3	1	3
$AF_{Intra\text{-}I=0.5\%}$	12.36	5.06	12.36	5.06
HD_M^I	1.4 mg/kg-d[a]	11.0[b]	0.34 mg/kg-d[a]	7.5[c]
	P05	**P95**	**P05**	**P95**
$HD_M^{I\,(d)}$	**0.12 mg/kg-d**	15 mg/kg-d	**0.04 mg/kg-d**	2.5 mg/kg-d

[a] HD_M^I (P50) = IPCS POD / ($AF_{Interspecies\text{-}BS}$ x $AF_{Interspecies\text{-}TK/TD}$ x $AF_{Intraspecies}$)
[b] (Composite P95/P50) = $10^{\wedge}[(\log 3.95)^2 + (\log 1.26)^2 + (\log 3)^2 + (\log 5.06)^2]^{0.5} = 11.0$
[c] (Composite P95/P50) = $10^{\wedge}[(\log 1.51)^2 + (\log 1.26)^2 + (\log 3)^2 + (\log 5.06)^2]^{0.5} = 7.5$
[d] HD_M^I (P05) = HD_M^I (P50) / (Composite P95/P50). HD_M^I (P95) = HD_M^I (P50) x (Composite P95/P50)

Calculation of HD_M^I from chronic DINP exposure: liver toxicity (Incidence = 0.1%)

Aspect	Spongiosis hepatis		Increased serum ALT	
	P50	P95/P50	P50	P95/P50
BMD	31.9 mg/kg-d	3.72	23.4 mg/kg-d	1.51
BMD-to-ED_{50} adjustment	3	1.5	N/A	N/A

IPCS POD = ED_{50}	**95.6 mg/kg-d**	**3.95[c]**	**23.4 mg/kg-d**	**1.51**
$AF_{Interspecies-BS}$	5.64	1.26	5.64	1.26
$AF_{Interspecies-TK/TD}$	1	3	1	3
$AF_{Intra-I=0.1\%}$	20.42	6.99	20.42	6.99
HD_M^I	0.8 mg/kg-d[a]	13.9[b]	0.2 mg/kg-d[a]	9.8[c]
	P05	**P95**	**P05**	**P95**
$HD_M^{I\,(d)}$	**0.06 mg/kg-d**	12 mg/kg-d	**0.02 mg/kg-d**	2.0 mg/kg-d

[a] HD_M^I (P50) = IPCS POD / ($AF_{Interspecies-BS}$ x $AF_{Interspecies-TK/TD}$ x $AF_{Intraspecies}$)
[b] (Composite P95/P50) = $10^{[(\log 3.95)^2 + (\log 1.26)^2 + (\log 3)^2 + (\log 6.99)^2]^{0.5}} = 13.9$
[c] (Composite P95/P50) = $10^{[(\log 1.51)^2 + (\log 1.26)^2 + (\log 3)^2 + (\log 6.99)^2]^{0.5}} = 9.8$
[d] HD_M^I (P05) = HD_M^I (P50) / (Composite P95/P50). HD_M^I (P95) = HD_M^I (P50) x (Composite P95/P50)

Calculation of HD_M^I from chronic DINP exposure: liver toxicity (Incidence = 0.01%)				
Aspect	**Spongiosis hepatis**		**Increased serum ALT**	
	P50	**P95/P50**	**P50**	**P95/P50**
BMD	31.9 mg/kg-d	3.72	23.4 mg/kg-d	1.51
BMD-to-ED_{50} adjustment	3	1.5	N/A	N/A
IPCS POD = ED_{50}	**95.6 mg/kg-d**	**3.95[c]**	**23.4 mg/kg-d**	**1.51**
$AF_{Interspecies-BS}$	5.64	1.26	5.64	1.26
$AF_{Interspecies-TK/TD}$	1	3	1	3
$AF_{Intra-I=0.01\%}$	37.71	10.39	37.71	10.39
HD_M^I	0.4 mg/kg-d[a]	18.9[b]	0.1 mg/kg-d[a]	13.9[c]
	P05	**P95**	**P05**	**P95**
$HD_M^{I\,(d)}$	**0.02 mg/kg-d**	8 mg/kg-d	**0.008 mg/kg-d**	1.5 mg/kg-d

[a] HD_M^I (P50) = IPCS POD / ($AF_{Interspecies-BS}$ x $AF_{Interspecies-TK/TD}$ x $AF_{Intraspecies}$)
[b] (Composite P95/P50) = $10^{[(\log 3.95)^2 + (\log 1.26)^2 + (\log 3)^2 + (\log 10.39)^2]^{0.5}} = 18.9$
[c] (Composite P95/P50) = $10^{[(\log 1.51)^2 + (\log 1.26)^2 + (\log 3)^2 + (\log 10.39)^2]^{0.5}} = 13.9$
[d] HD_M^I (P05) = HD_M^I (P50) / (Composite P95/P50). HD_M^I (P95) = HD_M^I (P50) x (Composite P95/P50)

Calculation of HD_M^I from chronic DINP exposure: liver toxicity (Incidence = 0.001%)				
Aspect	**Spongiosis hepatis**		**Increased serum ALT**	
	P50	**P95/P50**	**P50**	**P95/P50**
BMD	31.9 mg/kg-d	3.72	23.4 mg/kg-d	1.51
BMD-to-ED_{50} adjustment	3	1.5	N/A	N/A
IPCS POD = ED_{50}	**95.6 mg/kg-d**	**3.95[c]**	**23.4 mg/kg-d**	**1.51**
$AF_{Interspecies-BS}$	5.64	1.26	5.64	1.26
$AF_{Interspecies-TK/TD}$	1	3	1	3
$AF_{Intra-I=0.001\%}$	64.25	14.65	64.25	14.65
HD_M^I	0.3 mg/kg-d[a]	25.0[b]	0.06 mg/kg-d[a]	18.9[c]
	P05	**P95**	**P05**	**P95**
$HD_M^{I\,(d)}$	**0.01 mg/kg-d**	7 mg/kg-d	**0.003 mg/kg-d**	1.2 mg/kg-d

[a] HD_M^I (P50) = IPCS POD / ($AF_{Interspecies-BS}$ x $AF_{Interspecies-TK/TD}$ x $AF_{Intraspecies}$)
[b] (Composite P95/P50) = $10^{[(\log 3.95)^2 + (\log 1.26)^2 + (\log 3)^2 + (\log 14.65)^2]^{0.5}} = 25.0$
[c] (Composite P95/P50) = $10^{[(\log 1.51)^2 + (\log 1.26)^2 + (\log 3)^2 + (\log 14.65)^2]^{0.5}} = 18.9$
[d] HD_M^I (P05) = HD_M^I (P50) / (Composite P95/P50). HD_M^I (P95) = HD_M^I (P50) x (Composite P95/P50)

Interpretation of Results

The National Academies and the WHO/IPCS have both recommended using the lower confidence limit (LCL) on a probabilistic dose-response distribution for use in decision-making, in place of a traditional reference dose (RfD) or reference concentration (RfC). The National Academies said in *Science and Decisions* that:

> Multiple risk-specific doses could be provided…in the various risk characterizations that EPA produces to aid environmental decision-making.[13]
>
> A Risk-Specific Reference Dose: For quantal effects, the RfD can be defined to be the dose that corresponds to a particular risk specified to be de minimis (for example, 1 in 100,000) at a defined confidence level (for example, 95%) for the toxicity end point of concern.[14]

The WHO/IPCS said:

> The LCL of the HD_M^I can be used as a probabilistic RfD to replace the deterministic RfD. In this case, the probabilistic RfD is the dose that protects the population from a specified magnitude and incidence of effect with a pre-specified per cent coverage (confidence).[15]

Consistent with the guidance from the National Academies and the IPCS, we summarize the above results by focusing on the lower confidence limit (5th percentile or P05) risk-specific doses (HD_M^I) for multiple levels of risk (incidence or I).

Based on application of the WHO/IPCS methodology to DINP liver effects from chronic exposures, we find that:

0.44 mg/kg-day is the lower bound (95% confidence) chronic human dose at which liver lesions are expected in 5% of the exposed population, and 0.17 mg/kg-day is the lower bound (95% confidence) chronic human dose at which reduced serum ALT is expected in 5% of the exposed population

0.18 mg/kg-day is the lower bound (95% confidence) chronic human dose at which liver lesions are expected in 1% of the exposed population, and 0.065 mg/kg-day is the lower bound (95% confidence) chronic human dose at which reduced serum ALT is expected in 1% of the exposed population

0.12 mg/kg-day is the lower bound (95% confidence) chronic human dose at which liver lesions are expected in 0.5% of the exposed population, and 0.04 mg/kg-day is the lower bound (95% confidence) chronic human dose at which reduced serum ALT is expected in 0.5% of the exposed population

0.06 mg/kg-day is the lower bound (95% confidence) chronic human dose at which liver lesions are expected in 0.1% of the exposed population, and 0.02 mg/kg-day is the lower bound (95% confidence) chronic human dose at which reduced serum ALT is expected in 0.1% of the exposed population

0.02 mg/kg-day is the lower bound (95% confidence) chronic human dose at which liver lesions are expected in 0.01% (1-in-10,000) of the exposed population, and 0.008 mg/kg-day is the lower bound (95% confidence) chronic human dose at which reduced serum ALT is expected in 0.01% of the exposed population

0.01 mg/kg-day is the lower bound (95% confidence) chronic human dose at which liver lesions are expected in 0.001% (1-in-100,000) of the exposed population, and 0.003 mg/kg-day is the lower bound (95% confidence) chronic human dose at which reduced serum ALT is expected in 0.001% of the exposed population.

The implications of these risk values can be understood by comparison with the exposure levels considered by EPA to represent negligible risk. EPA's assessment uses a POD of 3.5 mg/kg-day (HED) and a benchmark MOE of 30,[16] meaning that EPA concludes "risk is not considered to be of concern and mitigation is not needed"[17] for any DINP exposure below 0.12 mg/kg-day (3.5 mg/kg-day / 30 = 0.12 mg/kg-day). Our analysis indicates that an exposure of 0.12 mg/kg-day is

equal to the lower-bound dose for the 0.5% (1-in-200) risk level for spongiosis hepatis lesions, and an exposure of 0.12 mg/kg-day is greater than the lower-bound dose for the 1% (1-in-100) risk level increased serum ALT. These risks far exceed EPA's usual target range of protection for carcinogenic risks of 1-in-10,000 to 1-in-1,000,000.[18]

The estimates of HD_M^I presented here were based entirely on input values and equations available from the WHO/IPCS methodology document and related publications, and from EPA's *Draft Non-cancer Human Health Hazard Assessment for Diisononyl Phthalate (DINP)*. An important caveat to these calculations is that the values used to represent human variability may be understated. The IPCS default human variability distribution is based on 37 data sets for human toxicokinetic variability and 34 data sets for human toxicodynamic variability. Most of these data sets were obtained from controlled human exposure studies of pharmaceuticals conducted in small samples of healthy adults, representing considerably less variability than found in the general population.[19,20,21] If human variability is underestimated, then the actual dose associated with each incidence level (e.g. I =1%, I = 0.1%) will be lower than the values obtained from this analysis – or in other words, risk at each dose will be underestimated.