Environment International 145 (2020) 105820

Contents lists available at ScienceDirect

Environment International

journal homepage: www.elsevier.com/locate/envint

Application of US EPA IRIS systematic review methods to the health effects of phthalates: Lessons learned and path forward

ARTICLE INFO

Handling Editor: Adrian Covaci

1. About the special issue

This special issue presents several systematic reviews of the potential human health effects associated with exposure to phthalates and addresses some of the considerations and challenges that were encountered over the course of performing these reviews. This editorial presents the views of the lead U.S. Environmental Protection Agency (EPA) researchers on the project and the Special Issue Editors with regard to lessons learned, implications of methods, and the path forward for systematic reviews to support human health assessment.

Phthalates comprise a class of alkyl diesters of phthalic acid and are used in a variety of consumer products including cosmetics, personal care, pharmaceuticals, medical devices, children's toys, food packaging, and cleaning and building materials. Owing to this widespread use, human exposure to phthalates is ubiquitous across all life stages, including during gestation and early postnatal life (CPSC, 2014; NRC, 2008). There is a considerable body of evidence demonstrating that certain phthalates can disrupt male reproductive development through the inhibition of testosterone production as well as through mechanisms that are independent of the antiandrogenic effects (Foster and Gray, 2013; CPSC, 2014; NRC, 2008); therefore, recent assessments of phthalate toxicity have generally centered on male reproductive effects. A 2017 systematic review of the low dose toxicity of phthalates by the National Academy of Sciences (NAS) focused exclusively on three androgen-dependent male reproductive endpoints (testosterone measurements, anogenital distance, and hypospadias) that are known to be sensitive to phthalate exposure (NAS, 2017). Antiandrogenic activity was also the basis for the US Consumer Product Safety Commission (CPSC) rule to permanently ban certain phthalates at any amount greater than 0.1% in children's toys and childcare articles, which was expanded in 2017 to include additional phthalates that share this mode of action (CPSC, 2017).

The work described in this Special Issue was designed to characterize the range of health outcomes associated with exposure to phthalates, including emerging health outcomes that were not covered by recent reports, and consists of six systematic review papers and three methodology papers. Systematic reviews of epidemiology studies were conducted for six phthalates (DEHP, DBP, BBP, DINP, DIBP, DEP) and are reported in four papers that describe associations of these chemicals with male reproductive effects (Radke et al., 2018), female reproductive and developmental effects (Radke et al., 2019a), metabolic effects (Radke et al., 2019b), and neurodevelopmental effects (Radke et al., 2020). Systematic reviews of experimental animal studies were conducted for DIBP (Yost et al., 2019) and DEP (Weaver et al., accepted for publication), with each paper describing the evidence for six major health outcome categories: male reproductive, female reproductive, developmental, liver, kidney, and cancer. The methodology papers address issues encountered during the course of conducting these reviews that will be of broad applicability to practitioners of systematic review in environmental health, including the evaluation of epidemiological studies using outcome-specific assessment criteria (Radke et al., 2019c) and the evaluation of animal studies for reporting quality, risk of bias, and sensitivity (Dishaw et al., accepted for publication). Lastly, Blessinger et al. (2020) present an ordinal dose–response model that addresses the challenges of modeling the collection of endpoints that characterize "phthalate syndrome" in experimental animal studies.

We believe this experience in applying systematic review methods to a large, complex evidence base will be informative for others in the field. Systematic review is being increasingly recognized by research and regulatory organizations as the gold standard for chemical risk assessment. Many of the benefits of systematic review, as well as the major advances and challenges facing researchers in this field, are illustrated in the July 2016 *Environment International* Special Issue, "Systematic Review Methods for Advancing Chemical Risk Assessment" (Whaley and Halsall 2016). Advantages of this approach include transparency and objectivity. However, challenges remain, including how to make systematic review pragmatic when applied to the broad topic areas encountered in environmental health. The methodologies for study selection, study evaluation, and evidence synthesis described in this set of papers were developed for use in the EPA's Integrated Risk Information System (IRIS) assessments to address some of these challenges, but these papers are among the first publicly available reviews using the methodologies. It is important to note that the IRIS approach described in the systematic review protocol for these reviews has evolved since their development and it is expected this evolution will continue over time as methodologies advance and feedback is received from the scientific community. Recently posted protocols reflect the current IRIS approach and are available at https://www.epa.gov/iris/iris-recent-additions.

https://doi.org/10.1016/j.envint.2020.105820
Received 14 May 2020; Accepted 17 May 2020

2. Challenges and lessons learned

There are several scientific issues that cut across multiple papers in this special issue, which will be discussed here. This includes the use of a novel study evaluation framework to evaluate reporting quality, risk of bias, and sensitivity, and the use of expert judgment in critical appraisal of individual studies (study evaluation) and the overall body of evidence. Specific considerations for epidemiology and experimental animal evidence of phthalate exposure are also described below.

2.1. Critical appraisal of evidence

All of the reviews of health effects of phthalates in human and animal studies (Radke et al., 2018, 2019a, 2019b, 2020; Yost et al. 2019; Weaver et al., accepted for publication) use the same general approach for study evaluation and evidence synthesis, though specific considerations vary between the disciplines. For each of the stages of the systematic review, the protocol (available in the supplemental materials for each review) describes the considerations that inform expert judgments on the evidence. For study evaluation, this consists of core and prompting questions, as well as specific considerations for reaching a rating for each evaluation domain (Tables 4 and 5 in the protocol). Greater detail on study evaluation is provided in publications in this special issue: Radke et al. (2019c) for epidemiology studies and Dishaw et al. (accepted for publication) for animal studies. For evidence synthesis, the primary considerations are adapted from Bradford Hill's considerations for causality (Tables 10 and 11 in the protocol). These considerations are then used together to reach an overall judgment of the evidence using a structured framework (Tables 12 and 13 in the protocol). This framework is conceptually similar to and is informed by the well-established Grading of Recommendations, Assessment, Development and Evaluations (GRADE) approach, but is designed to address challenges specific to the analysis of environmental health data rather than clinical evidence.

During peer review, there were some concerns that the criteria for reaching the different levels of confidence (within or across studies) were too reliant on expert judgment without clear constraints on that judgment. Without clear, consistent criteria, there is a risk of poor transparency for the decisions made in the systematic review. The approach used in these reviews was intentionally designed to rely on expert judgment rather than prescriptive criteria, pre-defined weights, or strict up/downgrading on the premise that experts are best equipped to make decisions on complex issues; however, but the peer review feedback highlighted the need to improve transparency and limit the potential for excess subjectivity. All of the review decisions are documented either within the publication or in linked materials (e.g., rationales for study evaluations are available in Health Assessment Workspace Collaborative (HAWC), a free and open source web-based software application, available at https://hawcprd.epa.gov/portal/). For all systematic reviews, there is a tension in the amount of constraint on expert judgment, and this will be worked on in refining this balance in future reviews.

Reviewers also commented on the unusual approach of separately evaluating reporting quality, risk of bias, and sensitivity in animal studies, which is different from the approach taken in other common study evaluation tools such as. the Navigation Guide, which rates neither reporting quality nor study sensitivity in assessing the evidence, and the National Toxicology Program's Office of Health Assessment and Translation (OHAT) risk of bias tool (NTP, 2015), which does not rate reporting quality. The reason the IRIS approach employs sensitivity as an assessment domain is that systematic review community has struggled with how to capture certain aspects of study design that do not strictly fall under risk of bias, which is defined as "a systematic error, or deviation from the truth, in results or inferences". For example, a study could have been conducted in a way that is bias-free but looked at an insensitive period of exposure. Some tools include these in the risk of bias metrics, but the IRIS approach attempts to be more explicit in labeling them as "sensitivity" related. The IRIS "reporting quality" domain looks at whether a study has reported sufficient details to conduct a risk of bias and sensitivity analysis; studies that do not report basic information such as species and test compound may be excluded. The present view is this approach improves the transparency of the evaluation.

2.2. Epidemiology evidence

There are several phthalate-specific considerations for evaluating epidemiology evidence in these reviews, and these are described below. All are relevant to the four systematic reviews of epidemiology data (Radke et al., 2018, 2019a, 2019b, 2020).

2.2.1. Phthalate-specific study evaluation criteria for exposure measurement

The majority of epidemiology studies of phthalates are based on measurement of phthalate metabolite levels. An important consideration in the evaluation of these biomarkers is the matrix or tissue from which they are measured. Phthalate metabolite concentration in urine is considered to be the best proxy of exposure from all sources (ingested, dermally absorbed, inhaled). One problem with measurement in blood and other tissues is the potential for contamination from outside sources (Calafat et al. 2015). Phthalate diesters present from exogenous contamination can be metabolized to the monoester metabolites by enzymes present in blood and other tissues, but not urine. For this reason, studies using biomarker measurements in samples other than urine for the monoester metabolites (e.g. MEHP) were excluded from the systematic reviews. For secondary metabolites of long-chain phthalates (i.e. DEHP and DINP), which are not influenced by this issue, studies were considered low confidence when samples other than urine were used.

Another consideration in the measurement of phthalate exposure is the short half-life of phthalate metabolites in the body (ranging from approximately 3 to < 24 h). Exposure can vary by time of day as well as over time. The short-term (1–12 weeks) reliability of metabolite measures, measured by the intraclass correlation coefficient is approximately 0.3–0.6 for the shorter chain metabolites (DBP, DIBP, BBP, and DEP) and 0.1–0.3 for the long-chain phthalates (DEHP and DINP) (Johns et al., 2016). Thus, particularly for the long-chain phthalates, the use of a single sample is expected to introduce nondifferential exposure misclassification. More samples are needed to measure exposure to the long-chain phthalates with sufficient precision, and this is reflected in the study evaluations. However, given that the nondifferential misclassification is likely to bias results towards the null, this issue should not reduce confidence in observed associations with health effects.

Additional discussion of these exposure measurement issues as well as specific criteria are described in the systematic review protocol, available as supplemental material to each of the systematic reviews.

2.2.2. Potential confounding across phthalates

Since different phthalates may be used in similar applications, potential confounding among phthalates is another important consideration. Several phthalates have moderate correlations with each other in human urine. When results are similar for two moderately (or higher) correlated phthalates in a study, confounding by other phthalate exposure should be considered as a possible explanation. An ideal study would have accounted for this in its analysis, but this is still relatively uncommon and not always possible due to collinearity and other analytic issues. Rather than rating each study with lower confidence due to this issue, it was considered it an area of uncertainty for all of the conclusions about associations between individual phthalates and health effects.

2.2.3. Coherence of effects across phthalates

In addition to drawing conclusions for each phthalate-outcome pair, there is interest in analyzing coherence of effects across phthalates. This would inform conclusions about the effects of the total mixture of phthalates to which people are exposed as well as individual phthalates (i.e., coherence across phthalates would increase confidence in the effect). Unfortunately, available data made it challenging to evaluate this coherence. It was expected that the most similar phthalate pairs (i.e., DEHP/DINP and DBP/DIBP) would also have the most coherent evidence of effects, and this was generally not the case. However, for both pairs, the database for one phthalate (DINP and DIBP) was less sensitive due to the number of studies and lower exposure levels, which could explain the weaker evidence. Because of this limitation, coherence was difficult to assess.

2.2.4. Prioritizing outcomes for systematic review

There are 25 outcomes included in the epidemiology systematic reviews, totaling 150 phthalate-outcome pairs. However, this does not include every outcome that has been studied in the literature (75+). Due to resource constraints, there was a need to be pragmatic in the outcomes undergoing full systematic review. In addition, even with unlimited resources, there are outcomes where either the limitations of the evidence are so severe or the evidence so inadequate that there is little value in pursuing a review. The paper on metabolic effects (Radke et al., 2019c) provides an example of a "screening level" review that identified such outcomes (obesity and renal effects). In other cases, outcomes were dropped from further review due to inadequate data (cancer, hepatic and respiratory effects). Lastly, systematic reviews were planned for some additional effects (immune and thyroid) but were not completed due to available resources. A new approach, systematic evidence mapping (described in Wolffe et al., 2019), is now being used in the IRIS program for this type of scoping/prioritization exercise and is discussed further below.

2.2.5. Outcome-specific study evaluation criteria

Another consequence of the number of outcomes in the database was the need for outcome-specific criteria for study evaluation. It was not possible for the epidemiologists performing the reviews to be experts in every outcome. Therefore, a process of subject matter expert consultation was undertaken, and this is described in Radke et al. (2019c). The criteria developed as a result of this process are chemical agnostic, and thus can be adapted for future systematic reviews.

2.3. Experimental animal evidence

Cross-cutting issues encountered in the systematic review of experimental animal evidence are described below.

2.3.1. Using PECO to define outcomes for systematic review

Literature screening and the selection of relevant studies was guided by a PECO (Populations, Exposures, Comparators, Outcomes) statement, which was included in the systematic review protocol. Whereas the PECO for human studies included any examination of human health effects associated with phthalate exposure, the PECO for animal studies specified that the systematic reviews would focus on six broad outcome categories: male reproductive, female reproductive, developmental, liver, kidney, and cancer. These six categories were identified in a preliminary review of the literature as being commonly associated with phthalate exposure and were selected as a pragmatic approach to focus the evaluation on outcomes that were likely to have more data available. Indeed, as described in Yost et al. (2019) and Weaver et al. (accepted for publication), the systematic reviews of animal studies identified very few studies that evaluated outcomes other than those included in the PECO, and if other outcomes had been included it is unlikely that there would have been enough data available to draw conclusions. Thus, although the process for selecting these six outcomes was relatively informal, it helped to streamline the systematic reviews of DIBP and DEP and likely had minimal impact on the overall conclusions. Nevertheless, this approach runs the risk of missing potentially important health outcomes, and multiple reviewers called for increased transparency regarding how outcomes were selected for inclusion in the PECO. It is therefore recommended that systematic evidence mapping be employed as a more formal approach for outcome selection in future systematic reviews. Systematic evidence mapping has emerged as a methodology for identifying the amount and type of evidence available to address a particular topic (e.g. Miake-Lye et al., 2016; Wolffe et al., 2019), and can be used as a step during problem formulation in order to identify health outcomes that warrant a full systematic review.

2.3.2. Chemical- and outcome-specific considerations for animal study evaluation

Few phthalate-specific considerations were necessary for the evaluation of animal studies; the concerns about biomarker measurements and confounding of effects across phthalates described for human studies are not applicable to animal studies, which (as specified in the PECO) exposed animals to singular phthalates at a nominal dose. However, as with human studies, it was often necessary to develop outcome-specific considerations for animal study evaluation. Dishaw et al. (accepted for publication) provides examples of specific considerations for the evaluation of male reproductive outcomes, which were among the major outcomes evaluated in the DIBP and DEP animal studies.

2.3.3. Integration of mechanistic evidence

The benefits and challenges of incorporating mechanistic evidence into a systematic review is an emerging area of interest in the field of risk assessment. Although a full evaluation of mechanistic data was not performed in the systematic reviews of DIBP and DEP, relevant mechanistic data was used to augment the qualitative synthesis. For instance, multiple studies that evaluated testosterone levels in rats or mice also reported on testicular expression of genes or proteins involved in cholesterol homeostasis and steroidogenesis. The inclusion of these data in the qualitative synthesis supports observations of the effects (or lack thereof) of these phthalates on testicular steroidogenesis and provides insight into the potential mechanism.

While it is beyond the scope of these systematic reviews, integration across human and animal lines of evidence for phthalates would benefit from a more thorough evaluation of mechanistic data. As stated in their review of the EPA's IRIS process, the National Academies of Sciences, Engineering, and Medicine (2014) stated that, "When human data are nonexistent, are mixed, or consistently show no association and an animal study finds a positive association, the importance of mechanistic data is increased."

3. Conclusions and path forward

The work in this special issue addresses two main research areas: the effects of phthalate exposure on health and the application of systematic review methods to broad, complex health assessments. Regarding the first, this is the largest and most comprehensive analysis of the health effects of phthalates published to date. Other reviews have focused more narrowly on certain health outcomes or used non-systematic methodology. Given the ubiquity of phthalate exposure in today's society, there is a critical need to understand what that exposure means for people across all life stages. It is clear from these reviews that male reproductive effects are a hazard of phthalate exposure (Radke et al., 2018; Yost et al., 2019), consistent with past reviews, with the interesting observation that exposure to DEP (not generally considered a male reproductive toxicant) may lead to effects on sperm independent of androgenic effects (Weaver et al., accepted for publication). The review of animal studies also supports previous observations that DIBP is a more potent developmental toxicant than DEP and highlighted

some potential female reproductive effects of DIBP exposure (Yost et al., 2019; Weaver et al., accepted for publication). There is also strong evidence of associations with preterm birth (Radke et al., 2019a) and diabetes and insulin resistance (Radke et al., 2019b), which has not been established previously. In addition, these reviews indicated some areas where the evidence is not as strong as may be commonly believed. The evidence for an association with neurodevelopment was largely inconsistent without a clear pattern of association (Radke et al., 2020), and there are important data gaps that need to be filled before drawing any causal conclusions.

A major limitation of the work described in this special issue is the lack of evidence integration across the evidence streams (i.e., human, animal, and mechanistic evidence). An integration approach is aimed at identifying how complementary but different evidence streams might be combined, yielding knowledge of the effects of an environmental exposure which might not be available if the evidence streams are analyzed in isolation from each other. For example, epidemiology and experimental animal studies complement each other in identifying health effects from exposures, as they each bring strengths that address key limitations in the other discipline (i.e., observational epidemiology studies provide evidence in humans, the species of interest, but do not allow for control of exposure; while experimental animal studies allow for precise and randomized control of exposure but effects in animals may not be directly relevant to humans). This is an important future step for the evidence of health effects resulting from phthalate exposure and could inform some of the limitations described in the systematic reviews, such as reducing concerns about confounding across phthalates in the epidemiology studies.

In addition, this was one of the first comprehensive attempts to use systematic review methods developed by the EPA's IRIS program for the assessment of environmental chemicals. As a process, it has not been without its challenges. Systematic review methods were originally developed for confirmatory mode (Nosek et al., 2018) research projects posing only one or two questions or hypotheses. This is because analyzing existing evidence is both time-consuming and conceptually complex; reliably making sense of what existing research is saying in answer to a question has historically been considered feasible only for small bodies of evidence. Regulatory agencies such as EPA, however, have broader information demands than can be serviced by a focused systematic review, with questions about multiple exposure scenarios, populations of concern, and outcomes of interest needing to be addressed. The need to address multiple questions, and therefore synthesize large bodies of evidence, presents a major challenge to implementation of evidence synthesis methods in regulatory scenarios while remaining simultaneously systematic and timely. The reviews in this Special Issue are an attempt by IRIS scientists to achieve exactly this.

While systematic methods produce the most reliable results when summarizing existing evidence, it is not necessarily the case that specifically conducting a systematic review is the most appropriate approach in any given scenario. As described above, in scenarios where the primary need is to scope the literature, systematic evidence mapping methods may be the more fit-for-purpose approach (James et al., 2016). Evidence mapping can also be used to set priorities when a systematic review is required (Wolffe et al., 2019). For instance, in a regulatory scenario where there is suggestive evidence that a current exposure limit for a chemical may need to be revised, evidence mapping could be used to select the endpoints that are most likely to lead to a revision of the exposure limit, and full systematic review can then be conducted on those endpoints.

Where complex systematic reviews are deemed to be necessary, suitable research infrastructure needs to be developed which can support projects requiring detailed yet wide-ranging analysis that allows them to be conducted in appropriate time-scales with sufficient accuracy. Such infrastructure would include pre-trained and identified pools of experts, methodological templates and chemical-agnostic processes, and fit-for-purpose tools which can be applied reasonably quickly with sufficient validity – to name a few. These would also include artificial intelligence-supported systems and suitable data storage approaches facilitating rapid extraction and reuse of the detailed scientific information required by systematic assessment methods (Wolffe et al., accepted for publication; Wittwehr et al., 2020).

In conclusion, these reviews have improved our understanding of health effects of phthalate exposure, and lessons learned in the conduct of these reviews have informed the further development of systematic review methods in the IRIS program and may be relevant to human health risk assessments by other agencies. There are additional opportunities to improve acceptance of fit-for-purpose methods for this type of complex health assessment.

Declaration of Competing Interest

PW declares personal fees from Elsevier Ltd as Systematic Reviews Editor for *Environment International*, from the Cancer Prevention and Education Society, the Evidence Based Toxicology Collaboration, Yordas Group, and grants from Lancaster University, which are outside the submitted work but relate to the development and promotion of systematic review and other evidence-based methods in environmental health research, delivering training around these methods, and providing editorial services. The handling of this manuscript by *Environment International* complies with Elsevier policy on conflicts of interest. The other authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Blessinger, T.D., Euling, S.Y., Want, L., Hogan, K.A., Cai, C., Klinefelter, G., Saillenfait, A., 2020. Ordinal dose-response modeling approach for the phthalate syndrome. Environ. Int.

Calafat, A.M., Longnecker, M.P., Koch, H.M., Swan, S.H., Hauser, R., Goldman, L.R., Lanphear, B.P., Rudel, R.A., Engel, S.M., Teitelbaum, S.L., Whyatt, R.M., Wolffe, M.S., 2015. Optimal exposure biomarkers for nonpersistent chemicals in environmental epidemiology. Environ. Health Perspect. 123A, 166–168.

Consumer Product Safety Commission (CPSC), 2014. Chronic Hazard Advisory Panel on phthalates and phthalate alternatives. Workshop report. https://www.cpsc.gov/s3fs-public/qeenworkshopreport2016.pdf.

Consumer Product Safety Commission (CPSC), 2017. Prohibition of Children's Toys and Child Care Articles Containing Specified Phthalates: Determinations Regarding Certain Plastics. Federal Register, Docket No. CPSC-2016-0017.

Dishaw, L., Yost, E., Arzuaga, X., Luke, A., Kraft, A., Walker, T., Thayer, K., 2020. A novel study evaluation strategy in the systematic review of animal toxicology studies for human health assessments of environmental chemicals. Environ. Int. https://www.sciencedirect.com/science/article/pii/S0160412019335330?via%3Dihub (accepted for publication).

Foster, P., Gray, L.E., 2013. Toxic Response of the Reproductive System. Book Chapter in Casarett & Doull's Toxicology: The Basic Science of Poisons. McGraw-Hill Education, pp. 861–906.

James, K., Randall, N., Haddaway, N., 2016. A methodology for systematic mapping in environmental sciences. Environ. Evid. 5.

Johns, L.E., Cooper, G.S., Galiza, Meeker, J.D., 2016. Exposure assessment issues in epidemiology studies of phthalates. Environ. Int. 85, 27–39.

Miake-Lye, I.M., Hempel, S., Shanman, R., Shekelle, P.G., 2016. What is an evidence map? A systematic review of published evidence maps and their definitions, methods, and products. System. Rev. 5, 28.

National Academies of Sciences, Engineering, and Medicine (NAS), 2017. Application of Systematic Review Methods in an Overall Strategy for Evaluating Low-Dose Toxicity from Endocrine Active Chemicals. The National Academies Press https://doi.org/10.17226/24758.

National Research Council (NRC), 2008. Phthalates and Cumulative Risk Assessment: The Task Ahead. The National Academies Press https://doi.10.17226/12528.

National Toxicology Program (NTP), 2015. OHAT Risk of Bias Rating Tool for Human and Animal Studies. U.S Dept. of Health and Human Services, National Toxicology Program. https://ntp.niehs.nih.gov/ntp/ohat/pubs/riskofbiastool_508.pdf.

Nosek, B., Ebersole, C., DeHaven, A., Mellor, D., 2018. The preregistration revolution. PNAS 115, 2600–2606.

Radke, E.G., Braun, J.M., Meeker, J.D., Cooper, G.S., 2018. Phthalate exposure and male reproductive outcomes: a systematic review of the human epidemiological evidence. Environ. Int.

Radke, E.G., Glenn, B.G., Braun, J.M., Cooper, G.S., 2019a. Phthalate exposure and female reproductive and developmental outcomes: a systematic review of the human epidemiological evidence. Environ. Int.

Radke, E.G., Galizia, A., Thayer, K.A., Cooper, G.S., 2019b. Phthalate exposure and metabolic effects: a systematic review of the human epidemiological evidence. Environ. Int.

Radke, E.G., Glenn, B.G., Galizia, A.G., Persad, A., Nachman, R., Bateson, T., Wright, J.M., Navas-Ascien, A., Arroyave, W.D., Puett, R.C., Harville, E.W., Pollack, A.Z., Burns, J.S., Lynch, C.D., Sagiv, S.K., Stein, C., Cooper, G.S., 2019c. Development of outcome-specific criteria for study evaluation in systematic reviews of epidemiology studies. Environ. Int.

Radke, E.G., Braun, J.M., Nachman, R.M., Cooper, G.S., 2020. Phthalate exposure and neurodevelopment: a systematic review and meta-analysis of human epidemiological evidence. Environ. Int.

Weaver, J.A., Beverly, B.E.J., Keshava, N., Mudipalli, A., Arzuaga, X., Cai, C., Hotchkiss, A., Makris, S., Yost, E.E., 2020. Hazards of diethyl phthalate (DEP) exposure: a systematic review of animal toxicology studies. Environ. Int (accepted for publication).

Whaley, P., Halsall, C., 2016. Assuring high-quality evidence reviews for chemical risk assessment: Five lessons from guest editing the first environmental health journal special issue dedicated to systematic review. Environ. Int. 92–93, 553–555.

Wittwehr, C., Blomstedt, P., Gosling, J.P., Peltola, T., Raffael, B., Richarz, A., Sienkiewicz, M., Whaley, P., Worth, A., Whelan, M., 2020. Artificial Intelligence for chemical risk assessment. Comp. Tox. 13, 100114.

Wolffe, T., Whaley, P., Halsall, C., Rooney, A., Walker, V., 2019. Systematic evidence maps as a novel tool to support evidence-based decision-making in chemicals policy and risk management. Environ. Int. 130, 104871.

Wolffe, Taylor A.M., Vidler, John, Halsall, Crispin, Hunt, Neil, Whaley, Paul, 2020. A survey of systematic evidence mapping practice and the case for knowledge graphs in environmental health & toxicology. Toxicol. Sci. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7261145/ (accepted for publication).

Yost, E.E., Euling, S.Y., Weaver, J.A., Beverly, B.E.J., Keshava, N., Mudipalli, A., Arzuaga, X., Blessinger, T., Dishaw, L., Hotchkiss, A., Makris, S., 2019. Hazards of diisobutyl phthalate (DIBP) exposure: a systematic review of animal toxicology studies. Environ. Int.

Elizabeth G. Radke[a,*], Erin E. Yost[b], Nicolas Roth[c], Sheela Sathyanarayana[d], Paul Whaley[e]

[a] *US EPA Center for Public Health and Environmental Assessment, Washington, DC, United States*

[b] *US EPA Center for Public Health and Environmental Assessment, Research Triangle Park, NC, United States*

[c] *Swiss Centre for Applied Human Toxicology, University of Basel, Basel, Switzerland*

[d] *University of Washington and Seattle Children's Research Institute, Seattle, WA, United States*

[e] *Lancaster Environment Centre, Lancaster University, Lancaster LA1 4YQ, UK*

E-mail address: radke-farabaugh.elizabeth@epa.gov (E.G. Radke).

* Corresponding author.