Environmental Defense Fund

Comments on

DRAFT OHAT APPROACH FOR SYSTEMATIC REVIEW AND EVIDENCE INTEGRATION FOR LITERATURE-BASED HEALTH ASSESSMENTS

Submitted June 11, 2013

Introduction

We commend and support the National Toxicology Program's Office of Health and Assessment and Translation's (OHAT) efforts to apply systematic review procedures for conducting literature-based evaluations to assess the evidence that certain environmental factors cause adverse health effects. We believe that if done appropriately, the adoption of systematic review practices will promote transparency, objectivity, and efficiency in the development of OHAT literature-based health assessments that ultimately serve to protect public health.

Environmental Defense Fund (EDF) respectfully submits these comments on the Draft OHAT Approach for Systematic Review and Evidence Integration for Literature-Based Health Assessments[1] (draft approach) and by extension, the two case studies to which it has been applied—Systematic Review to Evaluate the Evidence for an Association between Perfluorooctanoic acid (PFOA) Or Perfluorooctane Sulfonate (PFOS) Exposure and Immunotoxicity (PFOA/PFOS case study),[2] and Draft Protocol for Systematic Review to Evaluate the Evidence for an Association between Bisphenol A (BPA) Exposure and Obesity (BPA case study).[3]

Comments on Specific Steps of the Proposal

I. *Step 1: Prepare Topic; and*
Step 2: Search for and Select Studies for Inclusion

We generally support the proposed processes for topic preparation and search and study selection. In particular, we appreciate the high degree of transparency afforded by the public documentation of study search terms and studies meeting the eligibility criteria. Similarly, we support the development of the study flow schematic that will help individuals understand and track which and why studies are found eligible or

[1] U.S. Department of Health and Human Services, National Toxicology Program. Office of Health Assessment and Translation. *Draft OHAT Approach for Systematic Review and Evidence Integration for Literature-Based Health Assessments- February 2013*. 26 February 2013. (Draft Approach)

[2] U.S. Department of Health and Human Services, National Toxicology Program. Office of Health Assessment and Translation. *Systematic Review to Evaluate the Evidence for an Association between Perfluorooctanoic acid (PFOA) Or Perfluorooctane Sulfonate (PFOS) Exposure and Immunotoxicity*. 9 April 2013. (PFOA/PFOS Case Study)

[3] U.S. Department of Health and Human Services, National Toxicology Program. Office of Health Assessment and Translation. *Draft Protocol for Systematic Review to Evaluate the Evidence for an Association Between Bisphenol A (BPA) Exposure and Obesity.* 9 April 2013. (BPA Case Study)

ineligible for further consideration. Discussed below are recommendations EDF believes will improve the execution of these two steps of the protocol.

Step 1: Prepare Topic

The stated objective of both case studies is to develop hazard identification conclusions ("known," "presumed," "suspected," or "not classifiable") that exposure to the chemical of focus is associated with specific outcomes of interest. As described in Step 7 of the case studies, "not classifiable" is meant to encompass both conclusions of "not classifiable" or "not identified to be a hazard to humans." [4] These two conclusions are entirely different and we recommend that OHAT distinguish and separate upfront in Step 1, and finally in Step 7, conclusions of "not classifiable" and "not identified to be a hazard to humans." In effect, "not identified to be a hazard to humans" should be added as a fifth potential hazard identification conclusion reached in a systematic review.

The foundation of a systematic review is set at Step 1 of the protocol where the scope and focus of the topic to answer the specific question(s) of the evaluation is determined. As such, it is critical that details provided at this point of the review be clear and accurate. Three of the specific aims in the PFOA/PFOS case study refer to the examination of "PFOA and PFOS" exposures and effects, when in fact the review is examining the immunological effects of PFOA and PFOS separately. As written, "PFOA and PFOS" implies the review is examining cumulative exposures and effects of these substances. Instances of "PFOA and PFOS" should be replaced with "PFOA or PFOS" to accurately reflect the focus of the review.

We recommend additional clarification and consistency in describing the study eligibility criteria. Enumerated and described below are several areas of unexplained inconsistency between the two case studies that should be either aligned or explained:

1. Though referencing the same category of eligibility criteria, the subtitle "Types of human studies and model systems"[5] is used in the PFOA/PFOS case study whereas the subtitle "Types of participants and model systems"[6] is used in the BPA case study. "Types of human studies and model systems" is the more appropriate subtitle since types of studies involving human

[4] PFOA/PFOS Case Study pg. 54 and BPA Case Study pg. 59.

[5] PFOA/PFOS Case Study pg. 3.

[6] BPA Case Study pg. 3.

participants (e.g., clinical trials) are a subset of the universe of the types of human studies to be assessed in the systematic reviews.

2. The BPA case study includes eligibility restrictions for "types of studies," whereas there are no such restrictions in the PFOA/PFOS case study. Specifically, the BPA case study indicates that, "Only studies with a control or referent group will be included. Case studies, case reports, and ecological studies in humans will be excluded. Animal and *in vitro* studies without a concurrent control will be excluded."[7] The PFOA/PFOS case study simply states, "There are no restrictions based on study design."
3. The eligibility criteria for "Types of human studies and model systems"[8] differ between the case studies, especially with regard to non-animal and non-human studies.

It is not clear why there are differences in the eligibility criteria between the two case studies given the similarity in their objectives and specific aims. For this reason, we recommend that additional clarification be given for any differences in the case studies' eligibility criteria. If in practice the eligibility criteria between the case studies are the same, we recommend using consistent, standardized language to reflect this and prevent confusion. Consistency will be important for future systematic reviews.

Step 2: Search for and Select Studies for Inclusion

We support and commend the significant effort proposed in the draft approach to identify relevant studies, including grey literature, for the conduct of a systematic review. However, we strongly recommend inserting *a priori* stopping rules or pre-determined date restrictions[9] on study searches in systematic review protocols to balance the ambitious search effort with the need to progress to and complete the next steps of the systematic review; and to accomplish the overall goal of producing assessments that provide timely information to the public about the state of the science on the potential hazards of certain environmental compounds.

The cases studies indicate that the bibliographies of eligible studies will be scanned to identify additional potential studies for inclusion. EDF strongly recommends also

[7] Ibid.

[8] PFOA/PFOS Case study pg. 3 and BPA Case Study pg. 3.

[9] Currently, both case studies indicate that, "…databases will be searched from inception to the present" (BPA case study pg. 6 and PFOA/PFOS case study pg. 5). "Present" is ambiguous and needs to be more specifically defined (e.g., a firm date or a certain point in the develelopment of a systematic review).

searching for studies that *cite* identified eligible studies. This additional search is important because it targets more recent literature, which will not be included in the bibliographies of identified eligible studies.

The process described for settling disagreements on a particular study's eligibility differs between the case studies. The PFOA/PFOS case study indicates, "Disagreements between the 2 screeners will be resolved by discussion, involving a third member of the review team if necessary."[10] The BPA case study indicates, "Disagreements between the 2 screeners will be resolved by each screener independently reviewing the conflicts noted in DistillerSR, modifying and discussing responses as appropriate to resolve, and arbitration by a third member of the review team if necessary."[11] The process for resolving study eligibility disagreements between reviewers should be the same for all systematic reviews and EDF recommends that consistent language be used to reflect this. If for some reason the arbitration process for study eligibility must deviate from normal protocol it should be noted how and why an alternative process is to be used. Finally, as with screening studies for eligibility, a process needs to be in place for settling disagreements between reviewers on the acceptability of additional references (i.e., references from the bibliographies of identified eligible studies and references that *cite* eligible studies.)

It is logical and reasonable to require a baseline amount of data to proceed with conducting a systematic review. OHAT is proposing that at least three studies, meeting inclusion criteria must be available. It is important, however, that OHAT considers not only the quantity but also the quality—as described in Step 4—of available studies. Three high quality studies would likely be sufficient to proceed with a systematic review, but four low quality studies would be insufficient. Therefore, the protocol should be revised to include a more flexible standard for proceeding to the full systematic review.

Overall, the process outlined in the OHAT proposal for searching and selecting studies helps to ensure objectivity and facilitates the ability of reviewers to follow and provide comments on Step 2 of the systematic review. Addressing the issues raised above, however, will strengthen this step of the process.

[10] PFOA/PFOS Case Study pg. 8.

[11] BPA Case Study pg. 8.

II. *Step 3: Extract Data from Studies*

EDF strongly supports public access to the data extraction tables and the use of brackets to distinguish information that is inferred, converted, or estimated from measured data.

With respect to study data extraction elements, we recommend the following modifications:

1. In addition to extracting funding sources of the study itself, any other funding sources study authors report receiving should be documented.
2. When extracting conflict of interest (COI) information it should be noted whether there is no statement of COI, or whether authors report that there is no COI.
3. A data extraction field detailing whether any part of a study or analysis was contracted to an outside laboratory facility/company should be included under "methods" for all study types.
4. In addition to the number of animals assigned per group, the number of animals actually assessed in each group should be recorded, as this will affect important data metrics relevant to study quality (e.g., variance, statistical significance, power).
5. The data extraction table indicates that "Diet & Husbandry" information will be extracted from animal studies, yet all of the corresponding data elements listed relate to diet. The table should be modified to include the specific husbandry elements to be extracted (e.g., cage ventilation/setup, animals housed individually versus grouped, etc.).
6. The data extraction elements under "results" for animal studies include NOEL, LOEL, and statistical significance of other dose levels. This field should be expanded and left more open-ended to account for any other relevant dose-response estimates. We recommend modifying this field as follows "NOEL/NOAEL, LOEL/LOAEL, BMD/BMDL, statistical significance of other dose levels, and other estimates of risk."
7. Additional study method information should be extracted for *in vitro* studies:
    a. Assay characteristics (*e.g.*, name of assay kit, dynamic range of assay, assay target).[12]

[12] We recommend OHAT look to the data use guidelines discussed in the DRAFT EPA SAB EHHC report on Advancing the Application of CompTox Research for Human Health Risk Assessment (1/29/13) for additional *in vitro* assay characteristics to extract: http://1.usa.gov/V6Ywhh.

b. If conducted by a biotech company, the name of the biotech company and the corresponding technology should also be extracted.

8. For as many data extraction fields as possible, the customizable DistillerSR data extraction forms should allow for multiple-choice input options including "other" with a free-text answer to allow for documentation of additional details.

EDF recommends that OHAT consider involving two members of the review team to perform the data extraction step. This will require the development of a process to settle data extraction disagreements between reviewers. While potentially resource intensive, two reviewers performing this step would minimize data extraction errors and improve the overall quality of the process.

We recommend that additional clarity be provided for the treatment of multiple publications using the same study. The OHAT proposal indicates that with regard to publications of the same study it,

> "…will include all reports but select a study to use as the primary report and consider the others as secondary publications. The primary report will generally be the publication with the longest follow-up, or for studies with equivalent follow-up periods, we will select the study with the largest number of cases or the most recent publication as the primary report. We will include relevant data from all reports, but if the same outcome is reported in more than one report we will use data from the primary report."[13]

The significance of the primary report designation for later steps in the hazard identification process is not clear, especially if all "relevant data from all reports" will be included in the systematic review. We recommend OHAT provide additional clarification regarding the influence of the primary versus secondary report designation. To the extent a primary report carries more weight in the hazard identification process; we believe that selecting a primary report solely on the basis of it having the longest follow-up is not the most prudent approach. Additional considerations, such as study quality and whether the outcome measured is primary or secondary, should also inform the selection of the primary report. Nevertheless, EDF strongly recommends that all relevant data from reports of the same study be included in the systematic review. Again, the text quoted above is confusing and would benefit from additional explanation.

[13] PFOA/PFOS Case Study pg. 10. BPA Case Study pg. 11.

III. *Step 4: Assessing the Quality of Studies*

EDF generally supports the risk of bias approach outlined by OHAT to evaluate study quality. In particular we support,

1. not assigning composite risk of bias scores for studies, which ultimately masks sources and patterns of bias, and impairs later steps in the systematic review approach (including the tiering of studies based on quality and consequently the risk of bias evaluation for the overall body of evidence);
2. the default of "probably high risk of bias" when reviewers are unable to obtain sufficient information to evaluate a risk of bias question—particularly a key risk of bias element question; and
3. retaining 1) risk of bias elements for which there is not yet sufficient empirical evidence regarding their influence on bias and 2) risk of bias elements that are not currently standardly reported. Retaining these elements will encourage researchers to include such information in future publications, provide the empirical data to determine whether or not these elements are useful in assessing study bias, and ultimately improve the systematic review process.

However, we are concerned that the OHAT proposal does not include conflict of interest as a risk of bias element, despite evidence from the clinical field that conflict of interest, in particular financial conflict of interest, does bias study results.[14] AHRQ recommends, with regard to financial conflict of interest in evaluating risk of bias:

> "(1) at a minimum, EPCs [Evidence-based Practice Centers] should routinely report the source of each study's funding; (2) EPCs should consider issues of selective outcome reporting at the individual study level and for the body of evidence; and (3) EPCs should conduct sensitivity analyses for the body of evidence when they have reason to suspect that the source of funding or disclosed conflict of interest is influencing studies' results."[15]

We concur with these three AHRQ recommendations, which are not fully incorporated into OHAT draft approach. The OHAT proposal addresses the first

[14] Viswanathan M, Ansari MT, Berkman ND, Chang S, Hartling L, McPheeters LM, Santaguida PL, Shamliyan T, Singh K, Tsertsvadze A, Treadwell JR. Assessing the Risk of Bias of Individual Studies in Systematic Reviews of Health Care Interventions. Agency for Healthcare Research and Quality Methods Guide for Comparative Effectiveness Reviews. March 2012. AHRQ Publication No. 12-EHC047-EF. Available at: www.effectivehealthcare.ahrq.gov

[15] Ibid.

recommendation of collecting information on funding source, but it does not explicitly consider funding source when evaluating selective reporting; there are no funding related questions in the accompanying guidance for assessing risk of bias appendices for either of the case studies. Furthermore, it is not clear whether OHAT will routinely conduct sensitivity analyses to assess the impact of disclosed conflict of interest or not. In some places the case studies indicate that OHAT *may,*[16] while in other places the case studies indicate that OHAT *will,*[17] stratify and assess the potential impact of funding source. EDF recommends that financial conflict of interest be an explicit risk of bias element—as is done in other systematic review protocols assessing the evidence for a relationship between an environmental exposure and an outcome [18]—and at a minimum that funding source sensitivity analyses be conducted for every systematic review to understand whether and how funding source impacts overall hazard identification conclusions.

Certain aspects of the study quality evaluation process are unclear and need further explanation. The case studies indicate that risk of bias is assessed at the outcome level;[19] however, it is not clear how outcome level risk of bias then is used to tier studies by quality—especially when a study may have more than one reported outcome.[20] OHAT should provide additional clarification on the relationship between outcome-level risk of bias and individual study quality as these evaluations inform the overall rating of the body of evidence.

Finally, EDF supports and encourages OHAT to pursue near-term research to develop a risk of bias tool specific to *in vitro* studies that can be used in future assessments. It is important that every single study used in any way to inform a systematic review be evaluated for risk of bias. In order to rate confidence in specific studies and across a body of evidence in a systematic manner, a tool is needed to assess the extent to which findings can be explained by other factors, including both confounding and bias. We agree with OHAT that the Toxicological data Reliability Assessment Tool (ToxRTool) focuses primarily on assessing reporting quality, and that reporting

[16] PFOA/PFOS Case Study pg. 19 and BPA Case Study pg. 18.

[17] PFOA/PFOS Case Study pg. 46 and BPA Case Study pg. 52.

[18] An Evidence-Based Medicine Methodology to Bridge The Gap Between Clinical and Environmental Health Sciences. Health Affairs May 2011. 30:5 931-937.

[19] PFOA/PFOS Case Study pg. 15 and BPA Case Study pg. 17.

[20] PFOA/PFOS Case Study pg. 19 and BPA Case Study pg. 22.

quality is not an appropriate metric for assessing the internal validity of *in vitro* studies.

IV. *Step 5: Rate Confidence in the Body of Evidence*

EDF supports the conceptual framework outlined to rate confidence in the body of evidence. The various tables used to capture and present factors increasing or decreasing confidence across multiple studies are clear, comprehensive, and transparent, and we encourage these tables and summaries be made publicly available online.

EDF recommends that OHAT provide additional clarification on the flow and groupings of studies as they proceed through various analyses in this step of the systematic review. A flow diagram, for example, could be particularly helpful.

Highlighted below are several areas that we found particularly confusing and needing more explanation:

1. "Most studies are xx tier" is the guidance criteria for determining whether to downgrade confidence in the body of evidence for risk of bias (see tables 12 and 13 of the case studies). "Most" is ambiguous and should be explicitly defined *a priori* in the protocol (e.g. 33%, 50%).
2. As described in Step 4, study quality tiers will be used to identify studies that are of high risk of bias on many elements for the purpose of potentially omitting them from additional consideration in Step 5,[21] and for informing overall judgments on quality of the data across the evidence base. However Step 5 of the case studies states that OHAT "*will* omit the 3rd tier risk of bias studies from consideration when determining confidence ratings,"[22] which is much more definitive. What is OHAT's intention here? Are there criteria that will be used to decide whether or not to omit 3rd tier risk of bias studies in determining confidence ratings? If so, what are these criteria?
3. How are directness scores for individual studies combined to determine whether and to what extent the confidence rating needs to be downgraded? It appears that *each* study is first evaluated for directness and potentially downgraded 1 or 2 levels depending on the number of directness issues. However, there is no explanation of how this translates into an overall

[21] PFOA/PFOS Case Study pg. 19 and BPA Case Study pg. 22.

[22] PFOA/PFOS Case Study pg. 37 and BPA Case Study pg. 42.

determination of whether to downgrade the confidence rating in the body of evidence.

4. Similarly to 3 above, it is not clear how an overall imprecision downgrade will be determined after evaluating imprecision across different study types and outcomes. Much more explanation is needed in Step 5 to describe how evaluations of individual or subsets of studies for each of the confidence rating factors are grouped and then combined to reach a single, overall upgrade or downgrade decision.

Risk of Bias Across Studies. In conducting a sensitivity analysis to assess the extent to which Tier 3 studies obscure findings from Tier 1 and 2 studies, OHAT should also evaluate differences in study features (e.g., study design, data analyses) to better understand if and how certain study features inform bias.

Unexplained Inconsistency. For the quantitative data synthesis, the draft protocol states that a meta-analysis will be considered if 3 or more unique studies with "sufficient study level and methodological homogeneity with respect to population or animal model, study design, study duration, dose or exposure level, and health outcome"[23] are found. What are the criteria that will be used to determine whether these are sufficiently similar? These criteria need to be determined and outlined *a priori* in the protocol.

We strongly recommend that OHAT evaluate the statistical power of studies to determine whether the inclusion of underpowered studies is driving any observed inconsistency. Unexplained inconsistency across studies results in a downgrading of the confidence rating. If underpowered studies are driving the inconsistency, then the inconsistency is no longer unexplained—having been caused by the underpowered studies—and therefore the confidence rating of findings should not be downgraded. The same logic and approach should be applied when inconsistency across studies is detected for other important aspects of study design (*e.g.*, timing and duration of time between exposure and outcome).

Directness and Applicability. We agree with OHAT that while levels and route of exposure are important for understanding study relevance at known human exposure levels, they should not be factored into an evaluation of directness for rating confidence across studies in a *hazard identification* review. As explained in the case studies, exposure considerations occur subsequent to reaching a hazard identification conclusion during a "level of concern" conclusion. If level or route of exposure is explicitly included in the question being addressed in a hazard identification review

[23] PFOA/PFOS Case Study pg. 40 and BPA Case Study pg. 44.

then that should be reflected in the inclusion and exclusion criteria for study eligibility.

Dose-response. Upgrading for consistent, non-monotonicity curves across studies should not require *a priori* evidence of non-monotonicity, as currently outlined in the OHAT proposal. Because these assessments attempt to integrate existing data in order to provide new insight, they should not be constrained by previous assumptions. We urge that this guideline be modified. EDF recommends an upgrade +1 for dose response when non-monotonicity is consistently observed in the evidence base, especially when the evidence base consists of high quality studies, even if this is not the presumed pattern.

Final Note. The method for rating the confidence in the body of evidence seems to rely heavily on a quantitative process that sums up the individual factors increasing or decreasing confidence. EDF would recommend against this step being strictly quantitative, because several of these characteristics may overlap and a direct summation could result in a double-counting of their effect on the confidence rating. After assessing and summing all individual factors, there should be a qualitative process to evaluate and summarize the effect of these factors together as a whole on upgrading and downgrading the body of evidence.

V. *Step 6: Translate Confidence Ratings into Level of Evidence for Health Effect*

EDF supports the approach outlined in Step 6 of the systematic review proposal. In particular, EDF appreciates the tabular representation of how considerations for causality are embedded within the different steps of the OHAT proposal. Though levels of evidence for health effects are described in terms of associations, the integration of considerations for causality (i.e., Bradford-Hill) throughout the process implies that the final assessment describe the strength of causal relationships.

We also recommend that OHAT maintain the following guidelines in determining the level of evidence for health effects: 1) a conclusion of "no health effect" can only be reached when there is high confidence in the body of evidence, and 2) a conclusion of "inadequate evidence" can only be reached when there is low or moderate confidence in the evidence

VI. *Step 7: Integrate Evidence to Develop Hazard Identification Conclusions*

Overall, we endorse the process used by OHAT to reach the hazard identification conclusion in the final step of the systematic review protocol. We support reaching a

conclusion of "known to be a hazard to human" if there is high level of confidence for health effects from human studies, regardless of the level of evidence in non-human studies. This is consistent with approaches used by U.S. National Toxicology Program to classify carcinogens[24] and EPA's 2005 Guidelines for Cancer Risk.[25]

We applaud OHAT's anticipation of the increased use of *in vitro* or mechanistic information to generate hazard data and the resulting need to develop a systematic review approach that is able to integrate these data streams. We support the use of strong, concordant, biologically related *in vitro* or mechanistic information to upgrade hazard identification conclusions. *In vitro* and mechanistic information can provide additional support and clarity to outcomes observed in human and animal studies.

However, we have concerns with respect to the use of *in vitro* or mechanistic information to downgrade hazard identification conclusions especially when human data otherwise indicate a certain level of evidence. The BPA case study includes an extensive list of eligible mechanistic data: "ex vivo, cellular, genomic, or mechanistic outcomes reported in eligible animal or human studies; and data from cell systems, computational toxicology, high throughput screening data, and in silico models on interactions with key receptors involved in regulating adipogenesis…."[26] Downgrading based on mechanistic or *in vitro* information assumes that the relevant biological pathways and pathology of the health outcome of interest are well understood and characterized. The biology underlying (*e.g.*, relevant pathways, timing and patterns of exposure, etc.) and pathology of most complex health conditions and diseases (*e.g.*, cancer, reproductive and neurological disorders, etc.) is not fully characterized and may actually differ between affected individuals. The PFOA/PFOS case study itself notes, "Although some health effects of PFOA and PFOS are dependent on peroxisome proliferator-activated receptor alpha (PPARα, which shows strong species differences that may affect the relevance of animal data for human health), immune effects reported in laboratory animals appear to be partially or wholly independent of PPARα."[27] Further, the lack of coverage of the full biological response landscape is an oft-noted challenge with the newer high-throughput *in vitro* computational toxicology testing tools. In short, one would need

[24] "Report on Carcinogens Listing Criteria." *Department of Health and Human Services, National Toxicology Program*. 15 June 2011. 30 May 2013. http://ntp.niehs.nih.gov/?objectid=03C9CE38-E5CD-EE56-D21B94351DBC8FC3

[25] "Guidelines for Carcinogen Risk Assessment." *Environmental Protection Agency*. March 2005. 30 May 2013. http://www.epa.gov/raf/publications/pdfs/CANCER_GUIDELINES_FINAL_3-25-05.PDF

[26] BPA Case Study pg 3.

[27] PFOA/PFOS Case Study pg 2.

to understand the network/interplay of all relevant biological pathways and pathology associated with a particular health effect to have the confidence to downgrade a hazard identification conclusion using other relevant data. In contrast, it reasonable to upgrade hazard identification conclusions using *in vitro* or mechanistic information that is concordant and supports evidence from animal and human studies. For this reason, EDF recommends that OHAT not use other relevant data to downgrade the hazard identification conclusion—a concept analogous to reaching a conclusion of "known to be hazardous to human" based on human data, regardless of animal data in cancer risk assessments.

Other General Comments and Conclusion

EDF commends and supports the effort by OHAT to bring established and accepted systematic review approaches from the clinical realm into the field of environmental health. We strongly urge that the finalized OHAT systematic review approach maintains the high level of transparency outlined in this draft approach, from publicly tracking the study selection steps to documenting any deviations from specific protocols.

EDF supports and encourages OHAT to conduct the several study stratification analyses described in the case studies, such as tiering studies by quality to assess overall risk of bias, and sorting data both within and across studies in assessing dose-response relationships. Performing these types of stratification analyses will enable OHAT to better understand how individual or subgroups of studies influence final hazard identification conclusions.

In addition, EDF strongly recommends that deadlines for completing each of the steps of the systematic review, including comment periods, be established during the development of the systematic review protocol.

Environmental Defense Fund appreciates the opportunity to submit these comments and looks forward to their consideration by OHAT.