

INTRODUCTION

- The development of toxicity values based on human data is often preferred when sufficient evidence is available since it avoids uncertainties associated with animal data relevance. However, epidemiological study designs vary in uncertainties and risk of bias (RoB) potential, and present challenges when synthesizing such datasets for identifying causal relationships as well as relying on such data to quantitatively characterize exposure-response relationships in risk assessment.
- Per the NASEM meeting description, triangulation "refers to the practice of integrating results to inform and strengthen causal inferences. It provides a framework for considering and utilizing as much information as possible to address a research question. Its objective is to integrate results from different approaches, recognizing that each approach may have different, unrelated sources of potential bias"
- The objective herein is to provide a case study demonstration of "triangulation" in which a structured approach – in concert with logic grounded in biological plausibility concepts - was used to triangulate epidemiological evidence within a stream (other epidemiological data) as well as with experimental animal data as part of a larger effort to develop a plausible range of RfD values for TCDD.

METHODS

- A subset of exercises within a larger series of assessments were utilized to provide a case study of triangulation using epidemiological data in risk assessment for TCDD (Wikoff et al. 2019, Wikoff et al. 2021).
- In the systematic review process, epidemiological data were identified and selected using an a priori process documented in a protocol.
- Studies were critically appraised for internal validity per the NTP-OHAT RoB tool (NTP-OHAT, 2019). Key RoB domains included: Inclusion of appropriate confounders, exposure characterization, and outcome reporting.
- Confidence in the body of evidence (across the evidence stream) was further conducted using guidance from NTP-OHAT. Overall study tiering was based on the following:
- Tier 1: Study rated as "definitely low" or "probably low" RoB for all four key elements AND have most other applicable items answered, "definitely low" or "probably low" RoB.
- Tier 2: Study neither meets tier 1 or 3 criteria. Tier 3: Study rated as "definitely high" or "probably high" RoB for the majority of key elements, of which at least one is "definitely high," AND have most other applicable items answered, "definitely high" or "probably high" RoB.
- Biological plausibility and coherence in dose response were also assessed as part of triangulation within and across streams.
- Systematic review elements (risk of bias attributes or body of evidence attributes) were used to characterize uncertainty in the risk assessment.

Application of Risk of Bias for Environmental Epidemiology Evidence Characterization and Integration in Support of Risk Assessment: A case study evaluating the relationship between exposure to dioxin-like compounds (DLCs) and sperm count

Jon Urban¹, Daniele Wikoff², and Laurie Haws¹

ToxStrategies, Inc., 1-Austin, TX; 2-Asheville, NC

RESULTS

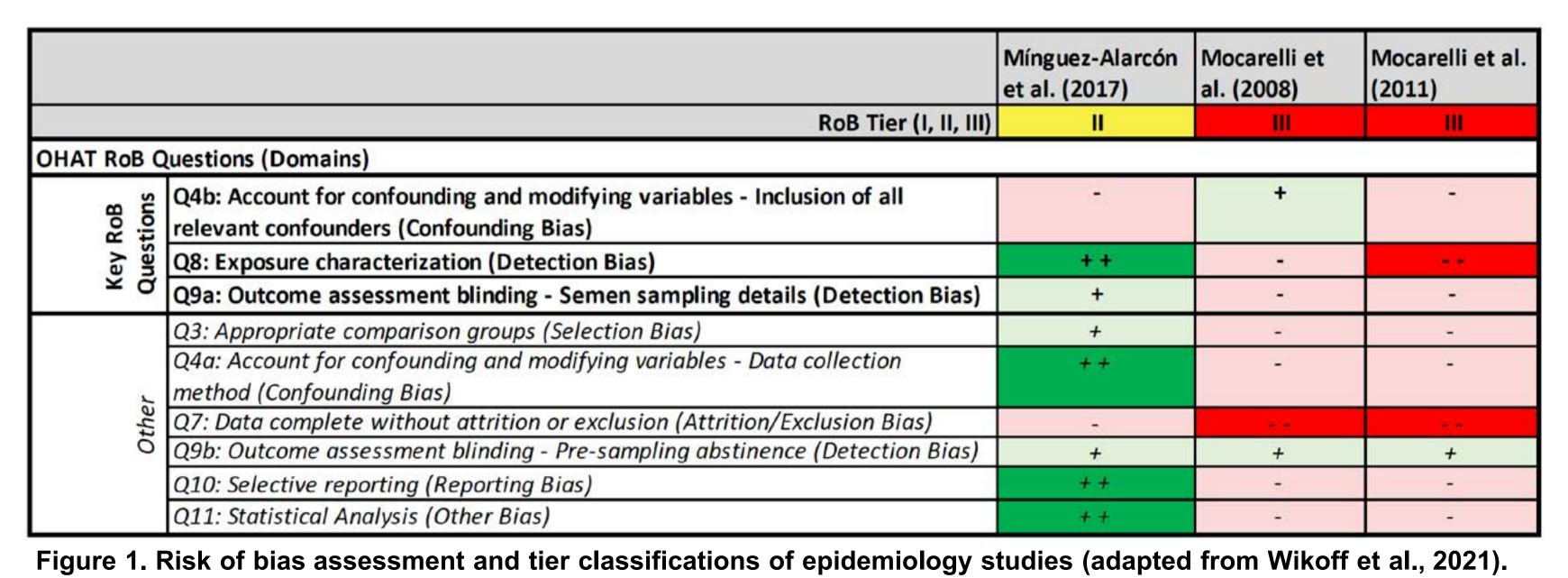
- The relevant epidemiological evidence base was limited to only three studies, from which study design information and dose-response data were extracted. (There were also 29 animal studies assessed as part of the larger assessment.)
- Study design and outcome information from the epidemiological studies are summarized in **Table 1** and include the following observations:
- Exposure to DLCs in two of the studies was based on an accidental exposure (an industrial explosion in Seveso, Italy in 1976). This resulted in high-level exposures to dioxins (primarily TCDD) over a short amount of time directly in
- The site of the other study other study was a town known for its proximity to chemical plant activities (Chapaevsk, Russia). Boys of families who grew up in this town were evaluated in young adulthood. The historical chemical manufacturing and processing activities near the town (chemical warfare agents, organochlorine pesticides, lead, dioxin-like compounds) make for a more complex exposure scenario.
- Although the populations in each of the three studies were limited in size, each of the three studies reported associations between TCDD and sperm quality endpoints to various degrees.

Table 1. Summarv of Human Studies that Investigated Association between Dioxin Exposure During Childhood and Effects n Sperm Endpoints as Adults (adapted from Wikoff et al., 2021)

Study Citation	Exposure Description	Study Population	Outcome	Dose-Response Analysis Reported by Authors?
Mocarelli et al. (2008)	Exposures to dioxins in Seveso, Italy resulting from industrial accident. Exposures defined by TCDD serum samples in boys collected at time of accident (1976) and at time of semen collection (1998), though only the earlier exposure measurements were included in sperm analysis. Other DLC PCDD/Fs not measured. No serum samples collected for control group; background exposures assumed to be ≤ 15 ppt in 1976.	Cohort - Subjects were 135 men exposed to TCDD while living in Seveso in Zone A in 3 age groups: infancy/prepuberty (1-9 years); puberty (10-17 years), and adulthood (18-26 years). Controls were 184 men of comparable ages who did not live in exposure area. Exposure/response measured in same subjects across 22-year time frame.	Significant reduction in mean sperm concentration (106/mL) between TCDD vs control groups for infant/prepuberty cohort; significant increase in mean total sperm count (106) between TCDD and control groups for older child/adolescent cohort; effects at individual exposure quartiles not tested. Ten highest exposure participants excluded from the authors' analysis. All reported sperm count data were above WHO's criterion for clinical significance (2010).	No - Dose-response data reported for serum TCDD quartiles in a figure, but no statistical analysis reported by authors based on quartiles. a
Mocarelli et al. (2011)	Exposures to dioxins in Seveso, Italy resulting from industrial accident (1976). Exposures used in sperm analysis were defined by TCDD maternal serum samples collected at time of accident (prior to pregnancy) and extrapolated to time of conception (between 1977-1984). Other DLC PCDD/Fs not measured. No serum samples collected for control group; background exposures assumed to be ≤ 15 ppt in 1976.	Cohort - Subjects were 39 men (21 breast-fed; mean age 22.5 years) born between 1977-1984 to mothers exposed to dioxin in Seveso accident (1976). Controls were 58 men born to mothers exposed to background dioxins. Exposure/response measured in different subjects (maternal exposure, response in sons).	Significant reduction in mean sperm concentration (106/mL) and mean total sperm count (106) between TCDD and control groups (all subjects and breastfed subjects, but not formula-fed subjects). All reported sperm count data were above WHO's criterion for clinical significance (2010).	Unclear - Statistical dose- response model cited by authors in results showing no statistical significance for semen concentration in the breast-fed men with the highest exposure, but model was not defined in study. No obvious trend analysis was reported for any sperm outcomes.
Mínguez-Alarcón et al. (2017)	Exposures in town of Chapaevsk, Russia. Environmental dioxins, metals and organochlorine pesticides associated with historical manufacture of chemical warfare agents, and industrial and agricultural chemicals. Exposures defined by DLC serum samples collected during enrollment (2003-2005).	Longitudinal Cohort - Subjects were 133 men originally recruited to Russian Boys Study between 2003-2005. No "unexposed" controls; subjects of low DLC serum levels used as comparators. Exposure/response measured in same subjects across 10-year timeframe (boys vs young men).	Significant reduction in median sperm concentration in two highest TCDD quartiles; significant reduction in sperm concentration (106/mL) and total sperm count (106) in three highest PCDD quartiles. Analyses were not adjusted for multiple coexposures previously reported for this population (metals, pesticides, etc.). All reported sperm count data were above WHO's criterion for clinical significance regarding fertility (2010).	Yes - Dose response data reported as predicted marginal means (95% confidence intervals) by quartiles of serum TCDD concentrations (represented as medians). Statistical dose-response trend analysis reported by authors. No significant differences or trends across PCDFs, DL-PCBs, or total TEQ.

Risk of Bias Assessment

• All three studies had high RoB for key domains (exposure, confounding, outcome), resulting in one Tier 2 study and two Tier 3 studies (Figure 1).



- All three studies had high RoB for key domains (exposure, confounding, outcome), resulting in one Tier 2 study and two Tier 3 studies (Figure 1).
- Applying NTP-OHAT data synthesis guidance, confidence in environmental epidemiological evidence stream for characterizing dose-response was categorized as low to very low.
- Per the NTP-OHAT confidence evaluation guidance, the human evidence stream would have started with a "low-to-moderate" level of confidence based on the features of the studies (lacked controlled exposure and had limitations in the control populations).
- Significant limitations in risk of bias for outcome assessment exist across the three studies.
- Insensitive endpoint: Sperm count and/or concentration reported in the human studies are highly variable parameters that are known to fluctuate dramatically over time (WHO, 2010).
- The low sample size, use of single samples, and inconsistency in time period between samples in the individual human studies likely contribute to explanation of inconsistencies. The large variations in measurement also decreased confidence related to low precision.
- The generalizability of the exposures in the two Seveso studies involving accidental, high-dose, acute exposure to TCDD relative to environmental exposures to DLCs is limited.
- To more comprehensively triangulate epidemiological data, biological plausibility of Figure 3. Sperm concentrations and TCDD exposure during childhood (1-9 years of age). Studies demonstrate a lack of an adverse response and coherence in the dose response between studies were
- The plausibility of an adverse response was limited given that all sperm measurements, regardless of exposure level, were well-above WHO criteria for clinically significant levels associated with screening for adversity regarding fertility.

 The biological plausibility of the exposure response-relationship was investigated by comparing dose-response data across studies presented in Figures 2 (TCDD) and 3 (DLCs). Although the direction of TCDD findings was consistent at the individual study level, the responses were not biologically coherent.

- Minguez-Alarcon et al. (2017) reported statistically significant decreases in sperm counts for upper TCDD and PCDD quartiles, but there was no statistical difference in Total TEQ measures.
- Despite having serum concentrations of TCDD three orders of magnitude higher, sperm concentrations reported in Mocarelli et al. (2008) were similar or greater than those reported by Minguez-Alarcon et al. (2017), demonstrating a lack of biological coherence between epidemiological studies.
- Studies in animal studies were inconsistent (not shown); several studies showed a lack of response, and in studies where a response was observed, it was generally at very high doses (which would possibly support findings from epidemiological studies involving the high-dose exposures from an accidental exposure but not from environmental levels).
- Taken together, the lack of TCDD dose-response across studies (across three orders of magnitude exposure), as well as a lack of biologically coherent dose response on a TEQ basis reduced confidence in the evidence.

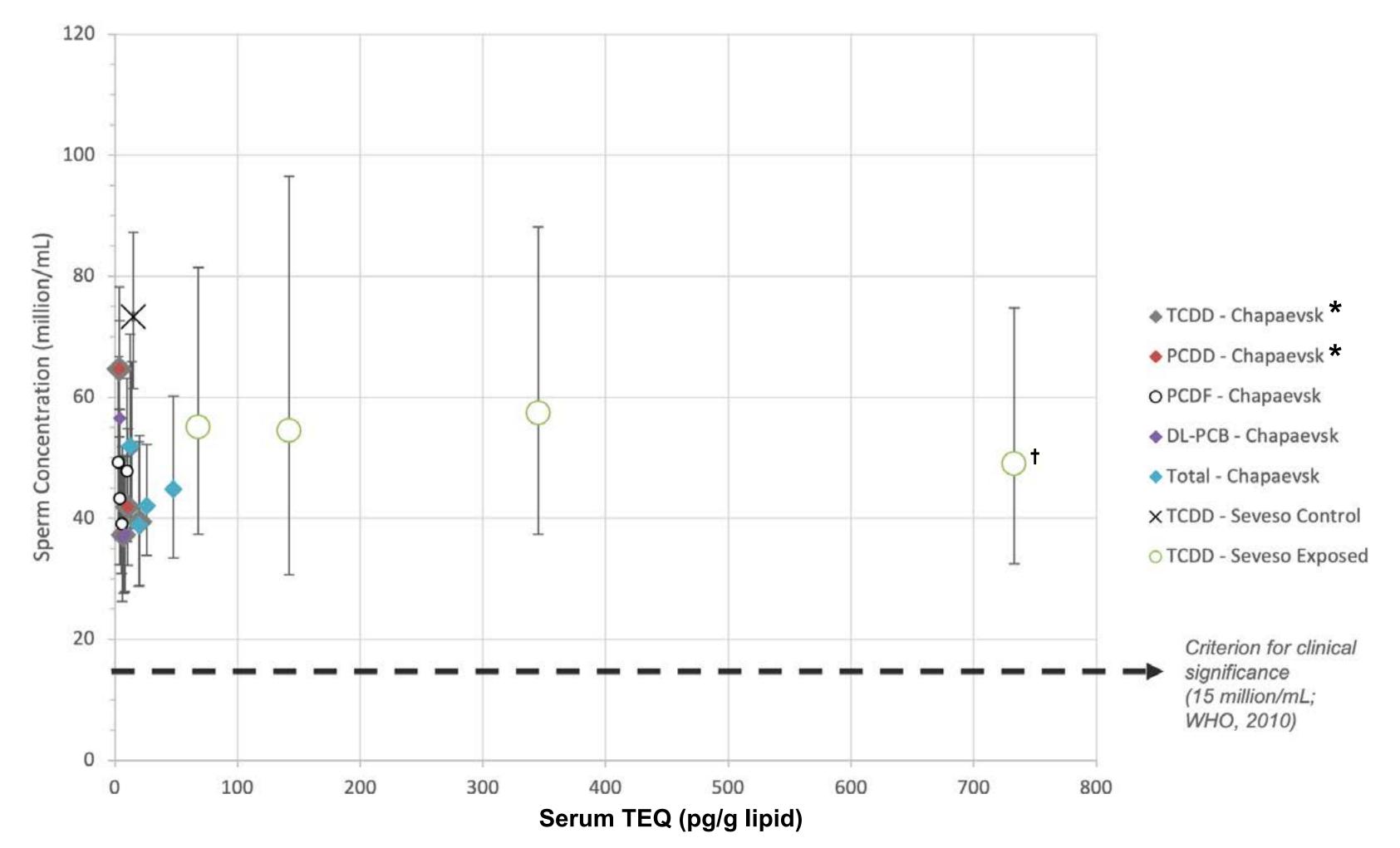
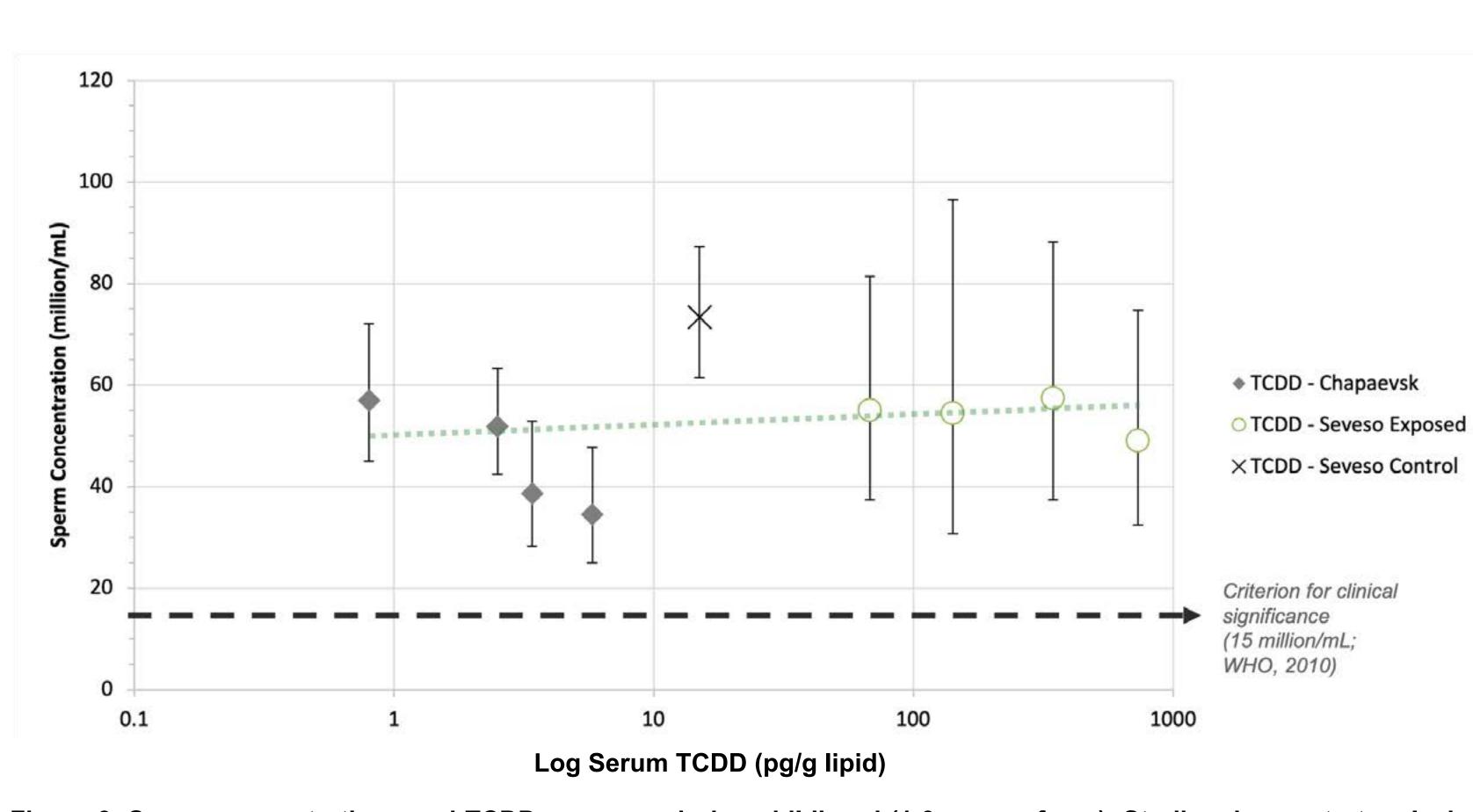


Figure 2. Sperm concentrations and DLC exposure during childhood (1-9 years of age). *Serum TCDD (3rd & 4th quartiles) & PCDD TEQ (2nd - 4th quartiles) statistically different than 1st quartile. †Mocarelli et al. (2008) did not statistically analyze Seveso quartiles vs. control. Wikoff et al. (2019) reanalysis showed only highest quartile (median of 733 TCDD pg/g lipid) was different than control (≤15 TCDD pg/g lipid) (adapted from Wikoff et al., 2021).



(Minguez-Alarcon et al., 2017) Seveso, Italy (Mocarelli et al., 2008) cohorts (adapted from Wikoff et al., 2021).

The systematic review techniques employed across assessment exercises also allows for characterization of uncertainty in the risk assessment.

- Using a "deconstruction" approach around the current EPA RfD, uncertainty can be qualitatively characterized, such as is demonstrated in Figure 4. Wikoff et al. (2019) also provides a demonstration of a quantitative uncertainty assessment based on key elements of risk of bias or body of evidence attributes.
- 2. If sufficient data are available, as was for TCDD, combined data analyses using meta-analytic techniques or Bayesian techniques, are useful tools to support causality assessment as well as risk assessment decisions. Such techniques are also demonstrated in Wikoff et al. (2019 and 2021).

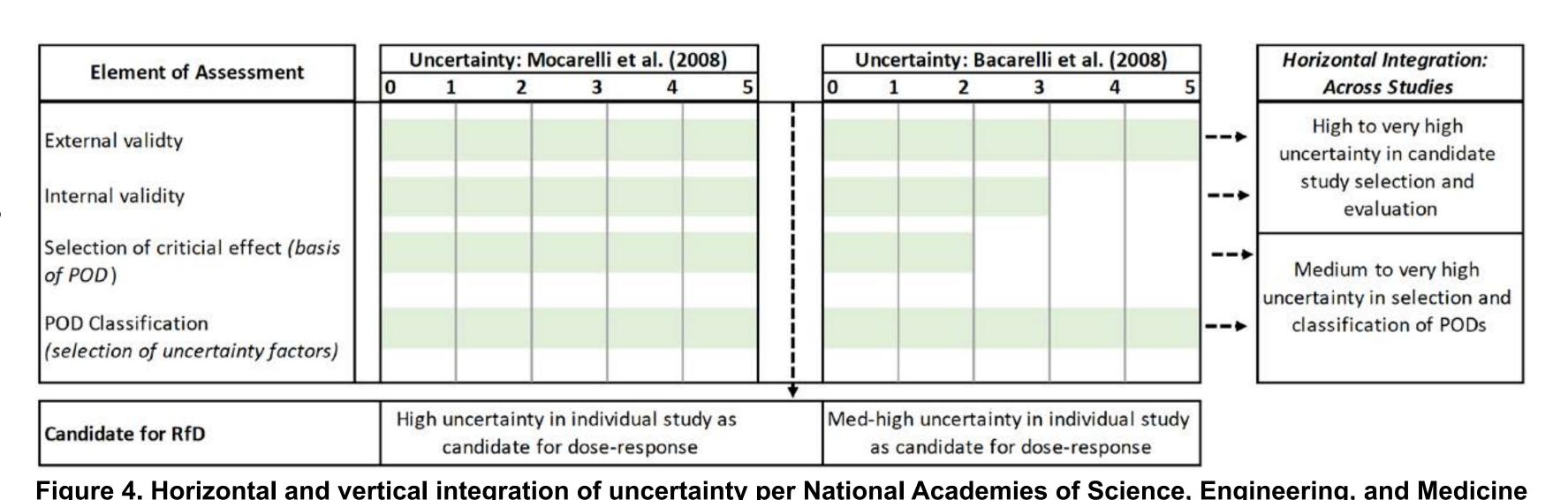


Figure 4. Horizontal and vertical integration of uncertainty per National Academies of Science, Engineering, and Medicine (2014) recommendations. POD, point of departure; RfD, reference dose (adapted from Wikoff et al., 2019)

CONCLUSION

The case study herein demonstrated:

- 1. The utility of the NTP-OHAT approach, when combined with biological plausibility and coherence considerations, in characterizing strengths and limitations when triangulating epidemiological evidence.
- 2. The importance of understanding the impacts of risk of bias within individual studies as well as biological plausibility and coherence concepts (including coherence in dose-response relationships) in developing both causality and risk assessment conclusions within and across evidence streams.
- The importance of logic in combining results of the evidence with considerations of the biology plausibility and coherence of the results in addition to simply assessing confidence in evidence in a structured triangulation of epidemiological evidence.
 - Standard GRADE-based approaches, for example, would not have readily facilitated the assessment of coherence in dose response or biological plausibility of an adverse response in the case study herein.

While systematic review brings many benefits, including a transparent structure for evaluation and triangulation, it is imperative that assessors not let the method for triangulation get in the way of logical risk assessment considerations, such as biological plausibility and dose concordance - within individual studies as well as across studies and evidence streams. This can be achieved by characterizing the *impact* risk of bias and other attributes commonly used to assess confidence rather than simply identifying such elements in the review process.

ACKNOWLEDGEMENTS AND DISCLOSURE

The case study demonstration of triangulation is based on a portion of a larger systematic review published by the authors in Wikoff et al. (2019; 2021).

ToxStrategies, Inc. is a consulting firm that provides services to private and public organizations on toxicology and risk assessment issues. As employees, the authors have provided prior consulting services to private stakeholders and public agencies on issues relevant to dioxin risk assessment as well as systematic review. No external funding was provided for the generation and presentation of this poster.

REFERENCES

Wikoff D, Haws L, Ring C, Budinsky R. (2019). Application of qualitative and quantita-tive uncertainty assessment tools in developing ranges of plausible toxicity values for 2,3,7,8tetrachlorodibenzo-p-dioxin. J Appl Toxicol. 39(9):1293-1310.

Wikoff DS, Urban JD, Ring C, Britt J, Fitch S, Budinsky R, Haws LC. (2021). Development of a Range of Plausible Noncancer Toxicity Values for 2,3,7,8-Tetrachlorodiben-zop-Dioxin Based on Effects on Sperm Count: Application of Systematic Review Meth-ods and Quantitative Integration of Dose Response Using Meta-Regression. Toxicol Sci. 179(2):162-182.

Other references available upon request.