National Academies of Science, Engineering, and Medicine (NASEM) Committee on EPA's Draft Formaldehyde Assessment: Excerpts of Key Critiques

August 9, 2023

While the executive summary and press release accompanying NASEM's is [being described as] largely positive (as NASEM's guide for committee chairs indicates, these summaries are largely the result of the chair and NASEM Study Director), NASEM's August 9 report contains substantive scientific and methodological criticism of EPA's draft assessment suggesting it is not the "best available science":

- "[T]he assessment does not satisfactorily follow recommendations for problem formulation and protocol development. EPA did not develop a set of specific protocols for the 2022 Draft Assessment in a fashion that would be consistent with the general state of practice that evolved during the prolonged period when the assessment was being developed…. The committee concluded that prepublished protocols are essential for future IRIS assessments to ensure transparency for systematic reviews in risk assessment." (pg. 5)
- "[F]or several noncancer outcomes, the committee was challenged to reconstruct the study evaluation approach and how the criteria were applied for study evaluation." (pg. 6)
- "The 2022 Draft Assessment reviewed by this committee was prepared across this decade of rapid change. Consequently, there was no static benchmark for evaluating the methods used. The committee took a broad view of the state of practice as it evaluated the 2022 Draft Assessment, recognizing that the methods used for that assessment would not correspond in all respects to the state of practice in 2023." (pg. xi)
- "In the discussion of uncertainties and confidence in the inhalation unit risk for myeloid leukemia, EPA should include the unknown dose rate-response relationship, the choice of statistical model and method, and the lack of understanding of mechanism. The three estimates in Table 2-35 should be presented as alternative, low-confidence inhalation unit risk estimates for myeloid leukemia without selection of a preferred estimate. EPA should not characterize the combining of other/unspecified leukemia with myeloid leukemia as 'the best approach.'" (pg. 10)
- "Many of the committee's findings and recommendations relate to the fragmented description of methods across the documents and insufficient clarity in their presentation.: (pg. 11)
- "[T]he committee's evaluation revealed some inconsistencies in how evaluation criteria are described and applied. Such inconsistency was broadly evident in the committee's review of EPA's evaluation of human and animal studies across noncancer outcomes (including for sensory irritation, pulmonary function, respiratory pathology, allergy and asthma, reproductive and developmental toxicity, and neurotoxicity)…. [T]he committee could not satisfactorily identify the final criteria that were applied, as well as the judgments made in determining overall study confidence for both human and animal studies. Inconsistencies between the stated criteria and the rationale for conclusions on study confidence were evident." (pg. 6-7)
- "[T]he committee found that the methods used for the assessment… reflect EPA's current practices in some components of he IRIS process." (pg. xi)
- "[T]here are opportunities to strengthen and clarify the 2022 Draft Assessment." (pg. xi)
- "In accordance with its statement of task, the committee did not conduct an independent hazard assessment to recommend alternative toxicity values." (pg. xi)
- "EPA provides criteria for study inclusion in the dose-response assessment, but does not include any discussion of how these criteria were applied to the specific studied chosen for dose-response." (pg. 8)
- "We also urge rigorous editing to enhance the overall quality of the assessment." (pg. xi)

- "The committee did not conduct an independent hazard evaluation or dose-response assessment, and therefore does not recommend alternative hazard identification conclusions or toxicity values. The committee also was not charged with commenting on other interpretations of scientific information relevant to the hazards and risks of formaldehyde, not did its statement of task call for a review of alternative opinions on EPA's formaldehyde assessment. Any other topics that do not fall within the committee's charge were beyond the purview of this study." (pg. 1)
- "In accordance with its statement of task, the committee did not conduct an independent assessment of formaldehyde's hazards or risks." (pg. 5)
- "The committee's review of the 2022 Draft Assessment documents the challenges faced by users of the assessment in navigating the voluminous documentation and understanding the methods used and evidence assessed. Revision is needed to ensure that the methods used for each outcome can easily be found." (pg. 5)
- "The origins of the various population, exposure, comparator, and outcome (PECO) statements are less clear. In particular, across noncancer outcomes, the rationale for excluding studies on the basis of the populations, exposures, and outcomes studies is not well documented." (pg. 6)
- "The committee's charge was to review the 2022 Draft Assessment prepared by EPA, and not to conduct its own formaldehyde assessment. The committee also was not charged with commenting on other interpretations of scientific information relevant to the hazards and risks of formaldehyde, or with reviewing alternative opinions of EPA's assessment. Any other topics not falling within the committee's charge were excluded from the committee's purview." (pg. 16)
- "Transparency in EPA's systematic review methods implies that the committee should be able to replicate each step based on the information included in the assessment documents or in publicly available supplemental materials. Accordingly, the committee used a case study approach to provide a detailed evaluation of the transparency and replicability of the 2022 Draft Assessment methods, relying on the documentation provided by EPA in the 2022 Draft Assessment and in the written responses to the committee's questions EPA's response to the committee's questions." (pg. 19)
- "This case study was conducted to test the replicability of EPA's approach to carrying out the eight steps of its IRIS assessment framework (Figure 2-2), as applied in the 2022 Draft Assessment. For this case study, the committee evaluated the various steps of the assessment for a single study and endpoint.… The committee carefully identified the criteria for study evaluation, finding inconsistencies across the several descriptions included in the 2022 Draft Assessment (Step 3). These inconsistencies complicated the committee's completion of Step 4—classification of study confidence by outcome. Overall, the committee found some inconsistencies between the presented human sensory irritation outcome– specific evaluation criteria and how they were applied to the Hanrahan et al. study, as well as how study limitations were presented for other sensory irritation studies…. With regard to the doseresponse analysis (Step 8), the committee had difficulty following EPA's reasoning in selecting the Hanrahan et al. study. This study was reported four decades ago as a two-page publication that does not meet the current norm for documentation and data access. The committee finds that the 2022 Draft Assessment does not adequately acknowledge the full scope of uncertainty associated with the Hanrahan et al. dose-response relationship." (pg. 28)
- "EPA should revise its assessment to ensure that users can find and follow the methods used in each step of the assessment for each health outcome. EPA should eliminate redundancies by providing a single presentation of the methods used in the hazard identification and dose-response processes." (pg. 28-29)

- "The terms used in the 2022 Draft Assessment—demonstrates, indicates, suggests, or is inadequate—are aligned with the IRIS Handbook, but are not obviously hierarchical, nor are they consistent with terms used elsewhere within EPA, (e.g., in the Integrated Science Assessments) or outside of EPA (e.g., the reports of the Surgeon Generals' on smoking and health). The terms also are used inappropriately, as in the inherent proposition that it is the evidence itself that demonstrates" or "indicates." The use of these terms represents an unnecessary source of inconsistency with the state of practice." (pg. 39)
- "EPA's inclusion of mechanistic evidence as a separate evidence stream is appropriate, but EPA faced substantial challenges in considering this evidence, especially in the context of systematic review." (pg. 39)
- "EPA drew on long-standing approaches for inferring causation….Nonetheless, the 2022 Draft Assessment deviates from those long-established approaches in several respects, including (1) blurring of the boundary between evidence synthesis and integration, and (2) the choice of terminology used to describe the strata in the four-level schema for classifying the strength of evidence." (pg. 40-41)
- "Despite the lack of evidence regarding systemic delivery of formaldehyde to distant sites, the biological basis for observed systemic effects (described in Chapters 4 and 5) remains unclear. Additional research is needed to address this apparent discrepancy." (pg. 48)
- "EPA should provide additional support for its decision not to use he BBDR model." (pg. 51)
- "EPA should address these shortcomings by updating tables and text to better document its dosimetry methods." (pg. 52)
- "EPA excluded outdoor exposure studies (Table A-32) without providing adequate justification. This exclusion was broadly applied to multiple health outcomes…. Excluding outdoor studies from the 2022 Draft Assessment may skew the evidence pool in the direction of higher exposure studies relative to the levels commonly experienced by the general population. The committee also found unconvincing EPA's argument that studies with lower exposure levels may have a limited ability to detect associations between formaldehyde exposure and health effects….EPA should include the body of evidence from outdoor exposure studies at the preliminary stage to derive a more holistic and objective assessment of the scientific literature." (pg. 57)
- "The outcome-specific criteria EPA used to evaluate the human studies were generally appropriate but the application of the specified criteria across studies appears inconsistent and it is not clear if the same set of quality criteria was applied uniformly across studies…. The committee could not find a consistent approach in how EPA evaluated the potential for selection bias, for example, across the range of observed response rates." (pg. 58)
- "EPA should document how it assessed the potential for different types of biases, the directionality of resulting biases, and the number of biases, and state how each combination should be interpreted in terms of high, medium, low, or not informative study confidence." (pg. 59)
- "EPA's approach appears to contradict the expert panel's advice that children's allergic symptoms are more liable to misdiagnosis than those of adults, particularly for asthma in infants and young children." (pg. 72)
- "[C]ombining all reproductive and developmental outcomes in a single group is an oversimplification." (pg. 76-77)
- "Information provided… for animal studies of male reproductive toxicity is inconsistent." (pg. 77)
- "I]t is not clear whether the study quality criteria were applied uniformly across studies." (pg. 80)

- "EPA should carefully address the following points regarding derivation of the RfC: Fully disclose data extracted from original study reports…. In reanalyzing data from published studies, the use of raw data is preferred… Avoid fitting a dose-response model that has as many parameters as the number of distinct aggregated data points taken from the published literature. Report and consider only models that meet the goodness-of-fit criteria EPA accepts…. Be more explicit as to how the final RfC was chosen." (pg. 83)
- "While the selection of this study is appropriate, and the major factors for its selection are described, the narrative discussion lacks clarity. The 2011 NRC committee called for EPA to 'develop, state, and systematically apply a set of selection criteria for studies and cancer end points' (NRC, 2011, p. 145)." (pg. 109)
- "EPA should acknowledge the uncertainty involved in interpreting the analyses on the degree to which exposure-response relationships are stronger than cumulative exposure for determination of peak exposure and risk." (pg. 111)
- "Insufficient information is given on the details of the dose-response analysis and the derivation of the unit risk values. For example, the 2022 Draft Assessment states, 'An adjustment was also made for the 15-year lag period" (p. 2-50), without explaining how the adjustment was made.'" (pg. 112)
- "The prediction of the number of annual incident cases in the U.S. at upper ends of outdoor (5 ppb) and indoor (20 ppb) formaldehyde exposure levels as a 'reality check' on the inhalation unit risk is a useful exercise that would be improved by acknowledging some of the possible environmental and other causative factors of nasopharyngeal cancer in the United States." (pg. 113)
- "EPA should provide additional detail on the modeling, including constraints imposed on model parameters, the results of model fitting (goodness-of-fit test), and the approach used o define lag parameters." (pg. 115-116)
- "EPA should address technical errors, such as mischaracterization of a trend p-value, with a thorough and technical edit and proofreading." (pg. 118)
- "[T] p-value is indicative of the lack of statistical significance for the association between myeloid leukemia death risk and cumulative formaldehyde exposure." (pg. 119)
- "other important sources of uncertainty include a lack of mechanistic support for myeloid leukemia, uncertainty about the true but unknown shape of the dose-response relationship and its data manifestation, and exposure misclassification." (pg. 120)
- "In part because of the weak dose-response relationship for cumulative exposure and myeloid leukemia data in the NCI cohort, the likely significant underreporting of myeloid leukemia, and the uncertainty in the optimal exposure metric, EPA determined that the inhalation unit risk estimate for myeloid leukemia is of low confidence. The committee concurs with the decision not to carry the myeloid leukemia risk estimate forward into the overall inhalation unit risk estimate for formaldehyde. Recommendation 5.15 (Tier 2): In the discussion of uncertainties and confidence in the inhalation unit risk for myeloid leukemia, EPA should include the unknown dose rate-response relationship, the choice of statistical model and method, and the lack of understanding of mechanism. The three estimates in Table 2-35 should be presented as alternative, low-confidence inhalation unit risk estimates for myeloid leukemia without selection of a preferred estimate. EPA should not characterize the combining of other/unspecified leukemia with myeloid leukemia as 'the best approach.'" (pg. 120)

- "A major uncertainty is the inability to include myeloid leukemia in the unit risk estimate because of the quality of the data available for dose-response analysis… the committee agrees that there is substantial uncertainty regarding extrapolation to lower doses…" (pg. 122)
- "[T]he committee recognizes that the IRIS Program turns to the older literature because of the exposure range and the potential to fit a model to estimate the dose-response relationship. This study was reported four decades ago as a two-page publication that does not meet the current norm for documentation and data access. Consequently, the agency takes work-around steps, including digitizing model estimates from a published figure that provides results of a logistic regression analysis. The committee finds that the full scope of uncertainty associated with the Hanrahan et al. dose-response relationship is not adequately acknowledged in the 2022 Draft Assessment." (pg. 148)
- "EPA needs to order the studies within industry by formaldehyde exposure levels, or the exposure difference between groups." (pg. 157)
- "Regarding sensory irritation… several studies have been identified as high or medium confidence, but only six were included in the sensory irritation dose-response analysis(Table 2-1). It is unclear why only these six studies were chosen." (pg. 158)
- "Figure 2-2 displays variability and uncertainty across all osRfCs along three dimensions: confidence, uncertainty, and risk size…. However, RPA was not explicit about how these factors were weighted." (pg. 160)
- "EPA proposed the overall RfC of 0.007 mg/m^3, but was not explicit about how it was chosen," (pg. 160)