(Date Received by EPA: January 5, 2022)

National Institute of Environmental Health Sciences

Comments on the Interagency Science Consultation Draft

EPA IRIS Toxicological Review of Formaldehyde – Inhalation (December 2021)

1. Assessment Development Methods and Organization

The Toxicological Review describes and applies a systematic review process for identifying, screening, and evaluating pertinent studies, and then for prioritizing the evidence to inform hazard and dose-response decisions. This process is described in the Toxicological Review's Preface on Assessment Methods and Organization, with documentation primarily in Appendix A.5. Please answer parts (a) and (b).

a) Please comment on whether the methods for assessment development (Preface on Assessment Methods and Organization) and the organization of the assessment are clear and transparent.

b) Please comment on whether there is sufficient documentation on methods and criteria for the following:

- *Identification of epidemiologic, experimental, and mechanistic studies (please identify any additional peer-reviewed studies that the assessment should consider).*
- *Critical evaluation of individual studies or sets of studies.*
- *Assessment of the weight of evidence (i.e., evidence integration).*
- *Selection of studies and data sets for deriving toxicity values.*

NIEHS Comments: EPA's IRIS program has advanced toward clearer and more transparent methods and standard approaches for the Toxicological Reviews. The review of these documents would be facilitated by readily providing access to the protocol with the review materials and making clearer both the methods used and where they have changed over time. The literature search and screening sections were the most challenging to follow. The study evaluations section, synthesis, and integration sections are better documented.

Tier 1

- The number of literature search terms for several health effect areas is minimal and a larger set of terms would be recommended. For example, A-344 PubMed "formaldehyde and (asthma or wheeze or respiratory or allergy or immune or sensitization) NOT ("formalin test" OR "formaldehyde fixation" OR "formalin fixation" OR "formalin fixed" OR "formaldehyde fixed" OR "formalin-induced" OR "formalin-evoked")".
  - The limited set of search terms is unlikely to capture all animal studies of hypersensitivity.
  - Evaluation considered eczema studies potentially informative; however, the literature search did not systematically search for those studies with the appropriate search terms.
  - Use of Not in a search string is potentially problematic. For example, a PubMed search would exclude potentially relevant studies that directly studied the appropriate endpoint (e.g., asthma) but also used formalin fixation.
- Interactive evidence maps are critical to an evaluation of the literature search and screening process. The links in Appendix F appear to require users to log in and access the assessment. This step was problematic for the review; therefore, the evidence mapping section was not fully evaluated.

Tier 2

- Main document pxxvLi26-31. The tracking of studies judged to have no impact on conclusions or toxicity values is done well in Appendix F; however, it is not clear as to how this was accomplished or why it was done. Since this is the functional methods section of the document, greater explanation is suggested, particularly in the absence of a protocol.
- Appendix pF-4Li 6-9. What is over-tagging and how was it ultimately resolved, or what impact did it have? Does it mean misclassification of studies?
- Screening - Identification of possibly impactful studies, Appendix F-5 li20-37. Insertion of a box with examples as to how the general considerations were applied to several Formaldehyde studies would be helpful.

- Appendix pF-4, Li 17-20. The logic as to formaldehyde's distribution following inhalation should be included in the problem formulation and PECO, and not in the screening section.
- Unless the search strategy changed, it is unnecessary to identify which literature search update studies were identified in the literature search diagrams (A-22, A-23, etc). It would be useful to specify whether they were identified in database searches or other sources and provide the basis for exclusions.

Tier 3

- It is challenging to get the right level of information in the overview, main document, and appendices. While the overview is shorter than the main text, it is a very long document. A suggestion is to reduce the text in the overview and be more explicit and detailed in the main document. This suggestion would apply for future reviews as well.
- A suggestion is to make the literature search figures (A-22, -23, A-24, etc.) interactive.

2. Toxicokinetics

Several assumptions and interpretations were applied in the Toxicological Review that were based on current research. Please answer parts (a), (b), and (c) considering the extent to which the available science on the toxicokinetics of inhaled formaldehyde is clearly presented and appropriately applied in the assessment of potential respiratory and systemic (i.e., nonrespiratory) health hazards.

a. Please comment on the Toxicological Review conclusion that inhaled formaldehyde is not likely to be distributed in appreciable amounts beyond the upper respiratory tract to distal tissues. This conclusion underpins the organization of the assessment and several key assumptions.

NIEHS Comments: Agree. This conclusion is supported, especially when inhaled (exogenous) formaldehyde and endogenous formaldehyde were evaluated and distinguished in the nasal and distal tissues (Lu et al 2010, 2011, 2012, and Moeller et al 2011).

- For example, Lu et al. (2010) evaluated formaldehyde induced DNA adducts and crosslinks. They were able to distinguish the difference between endogenously derived (^{12}C-formaldehyde) DNA adducts of formaldehyde from exogenously derived ($^{13}CD_2$-formaldehyde) DNA adducts in nasal epithelium following 1 or 5 days of exposure to 10 ppm inhaled $^{13}CD_2$-formaldehyde in F344 rats. However, in the same study no $^{13}C_2$-formaldehyde-DNA adducts (derived from exogenous source) were detected in the distal tissues (to the portal of entry) such as lung, spleen, liver, thymus, bone marrow or lymphocytes. In the same study, they also evaluated methanediol as a potential transport of inhaled formaldehyde; however, they did not see transitions to support methanediol derived from inhaled formaldehyde in the distal sites. Since the studies were in rats, it is important to note the species differences when extrapolating the data to humans, such as obligate nasal breathing in humans and oronasal breathing in humans as well as shape and complexity differences between rats and humans.
- Another study by Moeller et al (2010) evaluated the exogenous and endogenous formaldehyde in non-human primate bone marrow following exposure to inhaled $^{13}CD_2$-formaldehyde at 1.9 and 6.1 ppm exposure concentrations for 2 days (6hr/day). While the presence of endogenously derived DNA adducts (N^2-hydroxymethyl-dG) was confirmed, no exogenously derived DNA adducts were present. This confirms and further demonstrates the previous study findings (Lu et al. 2010) that inhaled formaldehyde is not likely to be distributed to distal tissues in appreciable amounts beyond nasal epithelium.
- In addition, the model developed by Corley et al (2015) reported that "formaldehyde does not penetrate deep into epithelial or subepithelial tissue even in the olfactory region where the penetration was greatest, and therefore does not transport directly to the systemic blood circulation at moderate exposure concentrations."'

b. Please comment on the Toxicological Review assumptions (based on [a]) that:

- *Inhaled formaldehyde is not directly interacting with tissues distal to the portal-of-entry (POE) to elicit systemic effects.*

NIEHS Comments: Agree. The earlier studies suggested that formaldehyde is not significantly absorbed into blood (Heck et al 1985, Casanova et al 1988, Kleinnijenhuis et al 2013). This conclusion is further supported by several studies, which were mentioned earlier (Lu et al 2010, 2011, 2012, Moeller et al 2011, Lai et al 2016, Yu et al 2015) and reported that adducts derived from exogenous formaldehyde were not detectable in distal tissues in experimental animals via inhalation route, suggesting that the exogenous formaldehyde is not systematically distributed. Also, studies showed that following inhalation of formaldehyde there was no covalent binding in bone marrow of rodents, supporting the lack of transport of formaldehyde to distal tissues.

- *Formaldehyde levels in the blood or at systemic sites are not changed as a result of formaldehyde from exogenous sources (inhalation).*

NIEHS Comments: Agree. It was previously hypothesized that inhaled formaldehyde might increase endogenous levels by methanediol releasing formaldehyde at distal sites. However, a vast number of references provide information about metabolic capacity in humans and animals that inhaled formaldehyde would not be expected to change/increase formaldehyde levels in blood (Heck and Casanova 2004, Casanova et al. 1988, Heck et al. 1985). In addition, the mathematical model developed by Franks et al. (2004) confirmed that if 1.9 ppm formaldehyde was inhaled, it was predicted to increase (0.00044 mg/L) blood concentrations well below measurable endogenous levels. Furthermore, studies following isotopically labelled formaldehyde inhalation did not significantly increase the endogenous formaldehyde levels (or adduct levels) when compared to control levels (e.g., Lu et al., Leng et al., or Moeller et al.).

- *Inhaled formaldehyde does not cause appreciable changes in normal metabolic processes associated with formaldehyde in distal tissues. Therefore, studies examining potential associations between levels of formaldehyde or formaldehyde byproducts in tissues distal to the POE (e.g., formate in blood or urine; brain formaldehyde levels) and health outcomes are not considered relevant to interpreting the human health hazards of inhaled formaldehyde.*

NIEHS Comments: Agree. Based on the studies and available data in the literature, it is unlikely that inhaled formaldehyde would be interacting with distal tissues to elicit systemic effect and as mentioned in the summary "inhalation exposure to formaldehyde has not been shown to cause significant changes to the tissue levels of formaldehyde in the nasal mucosa, the blood, or in the distal tissues". This might suggest that inhaled formaldehyde is not expected to alter the metabolic process in distal tissues, and therefore, the association between health outcomes in humans and formaldehyde products (and formaldehyde) in distal tissues would not be considered relevant for interpretation.

c. *Please comment on the Toxicological Review evaluation of the potential impact of normal levels of endogenous formaldehyde on the penetration and distribution of inhaled formaldehyde in the respiratory tract, on the basis of available dosimetric models and data.*

NIEHS Comments: While there are some dosimetric models (Schroeter et al. 2014, Campbell et al. 2020) and data from animal studies to evaluate the potential impact of endogenous formaldehyde levels on ADME of inhaled formaldehyde, additional studies addressing the potential impact of endogenous formaldehyde levels seems necessary to further evaluate the outcomes, especially when low exposure levels are considered. The two dosimetric models summarized in this review have limitations, such as endogenous levels calculated from blood concentrations (which doesn't correlate well with tissue concentrations), zero-order term (restriction on uptake form air phase to tissue compartment) being used for clearance, and one model is specifically for rats (which makes it hard, if not impossible, to use for extrapolation to humans). However, regardless of limitations, both models predicted the reduction of inhaled formaldehyde due to endogenous formaldehyde levels. Models, their limitations, and use of available data from the literature were well summarized in this review. Yet this highlights the need for additional studies and models to be used and to address the potential impact of endogenous formaldehyde on

ADME of inhaled formaldehyde in the respiratory tract, especially when evaluating lower levels of exogenous inhaled formaldehyde.

3. Respiratory System Health Effects (Noncancer).
For each noncancer POE health effect considered in the assessment and outlined in (a) to (e), below, please comment on whether the evidence integration decisions for hazard identification are clearly described and scientifically justified (considering the extent to which the available data have been appropriately synthesized to describe the strengths and limitations). In addition, please separately comment on whether the dose-response decisions are transparent and scientifically justified, including study selection for dose-response analyses; point of departure (POD) estimates, including modeling choices and assumptions, and dosimetric adjustments; selection of uncertainty factors and derivation of candidate values; selection of organ- or system-specific RfCs (osRfCs); and confidence in the calculated values. For these well-studied health effects, confidence was consistently judged as either medium or high.

Note: No comments on 3a (sensory irritation), b (pulmonary function), c (respiratory track pathology).

d. Allergy-related conditions

- *The assessment concludes that the available evidence indicates that formaldehyde inhalation likely causes increased allergic responses in humans, given the appropriate exposure circumstances. This conclusion was supported primarily by evidence in exposed humans, with supportive mechanistic evidence indicating that formaldehyde inhalation results in biological changes related to these outcomes in exposed animals.*
- *A POD from Annesi-Maesano et al. (2012), a human study, was ultimately selected to calculate an osRfC of 0.008 mg/m3 for allergy-related conditions. A UFC of 3 was applied to address UFH. This UFH value was selected using an evidence-based analysis.*

e. Prevalence of current asthma and degree of asthma control

- *The assessment concludes that the available evidence indicates that formaldehyde inhalation likely causes an increased frequency of current asthma symptoms or difficulty controlling asthma, given the appropriate exposure circumstances. This conclusion was supported primarily by evidence in exposed humans, with supportive mechanistic evidence indicating that formaldehyde inhalation results in biological changes related to these outcomes in exposed animals.*
- *PODs from the Annesi-Maesano et al. (2012), Krzyzanowski et al. (1990), and Venn et al. (2003) human studies were ultimately selected to calculate an osRfC of 0.006 mg/m3 for current asthma or degree of asthma control. A UFC of 3 or 10 was applied to address UFH. The UFH value applied to the POD from Annesi-Maesano et al. (2012) was selected using an evidence-based analysis.*

NIEHS Comments: The choice of studies used to contribute to the conclusions on allergy-related conditions, prevalence of current asthma and degree of asthma control, and the reasoning presented for those studies that were not selected is appropriate. Overall, the decisions are well described and scientifically justified. The human data support the conclusions that the evidence indicates that inhalation of formaldehyde likely increases the prevalence of allergic conditions, increases the prevalence of asthma symptoms, as well as decreases control of asthma symptoms in humans with moderate human data. The use of animal studies on asthma and respiratory allergy as used in the assessment, that is contributing to mechanistic information rather than to the conclusions on the likelihood of specific health effects, and to not use animal allergy or asthma data to derive reference values is appropriate. The selection of Annesi-Maesano et al., 2012 for both allergy-related and asthma was clearly described and presented. The selection of Matsunaga, Krayzanowski, and Venn was also transparently described as were the PODs from each study and the application of uncertainty factors.

Tier 1

- Overview p53. It is not clear why Liu et al., 2018 and Billonnet et al., 2001 studies are listed in the panel B of Figure 9 with the "High Exposures (>0.050 mg/m3)" when the midpoint of the Liu study is listed as

0.038 and Billonnet as 0.046. This requires correction as it does not seem to follow the legend's explanation of how studies were split by low and high exposure panels. For example, the high exposure group for Venn et al., 2003 and Matsunaga et al., 2008 are 0.046, which is the same as Billonnet and higher than Liu. If there is another reason, then further explanation should be provided in the text and legend. Note, the error is repeated in Figure 1-9 in the main text on page 1-93.

- Overview p55Li1-11, and Main document 1-120, and on … There is no mention of dose in the evaluation of mechanistic data. Is there any evidence that initial alterations occur at the same or lower dose than secondary alternation or effector-level changes? If not, then there is significantly less support for the mechanisms where higher concentrations are required for initial events than for later events. Similarly, the timing of events and the support or lack of support based on the available studies relative to the pathways in Figure 1-12 are not adequately addressed.

Tier 2

- Overview p47-48 and Main document p1-73Li3-4. Multiple sections indicate health effects that were not a focus of the review without providing an explanation. It is appropriate, helpful, and necessary for scientific judgments to be explained and this is generally done in the document. However, the statement that some endpoints were not the focus of the review without providing brief justification in the overview lessens the otherwise transparent use of the data. Recommend adding the reasoning where it has been omitted or pointing to the further details in the main document.
  - Overview p47Li26 and Main document p1-73Li3-4. "Dermal sensitization is not focus of this review" and p48Li1, "…infants and toddlers, but these outcomes were not the focus…"). The reader is left asking why. Suggest adding explanation throughout or in a brief separate section of the document to explain what is and what is not a focus of the review and the basis for making those decisions.
  - Overview p48Li10-11, Main document p1-73Li4-5. Statement is made that experimental animal studies were concluded to be unsuitable models without explanation. Suggest adding a brief reason to support the decision in the overview and at first mention in the main document. For example, animal models of airway hypersensitivity have limitations relative to human data for hazard characterization and therefore… See "Immune-mediated Conditions, Focusing on Allergies and Asthma, in Animal Studies" for more details.
  - Main document p1-73Li7-13, p1-74 line 14-19. Suggest adding brief text as to why the ability to respond to infection was not considered worth independent health hazard assessment rather than placing all the justification in the appendix.
  - Examples of appropriate, transparent, and helpful explanation: Overview p53Li3 –Lajoie et al., (2014) explains why the study was not included. As does the case definition text on p54Li7-9.
- Overview p55Li11 Figure 10, and Main document p1-121 Figure 1-12. Not explained why there are multiple relationship lines between some of the endpoints. In addition, some of the solid grey lines for "slight" appear to be solid while others appear dotted. Are there 2 types of "slight" relationships? This could be an artifact of the PDF and suggest using an alternate line width or formatting for clarity.
- Overview Table19 p58Li1. Table 19 would benefit from column headings.
- Main document p1-74Li24. Add the month of the literature search (May).
- Main document p1-80. Cannot easily read the overall confidence graphic for Annesi-Maesano 2012. This may be an issue with PDF, but the other studies are all clear.

4. Systemic (i.e., nonrespiratory) Health Effects (Noncancer).

For each noncancer systemic health effect considered in the assessment and outlined in (a) to (c), below, please comment on whether the evidence integration decisions for hazard identification are clearly described and scientifically justified (considering the extent to which the available data have been appropriately synthesized to describe the strengths and limitations). In addition, please separately comment on whether the dose-response decisions are transparent and scientifically justified, including study selection for dose-response analyses; POD estimates, including modeling choices and assumptions, and dosimetric adjustments; selection of uncertainty factors and derivation of candidate values; selection of osRfCs; and confidence in the calculated values. Confidence was consistently lower for these effects as compared with POE effects.

a. *Female reproductive or developmental toxicity:*
    - *The assessment concludes that the evidence indicates that inhalation of formaldehyde likely causes female reproductive or developmental toxicity, given the appropriate exposure circumstances. The conclusion for female reproductive or developmental toxicity is supported by evidence in humans, specifically, increases in time-to-pregnancy (TTP) and spontaneous abortion risk; mechanistic evidence explaining such effects without systemic distribution of formaldehyde is lacking.*

    - *A POD from Taskinen et al. (1999), a human study, was ultimately selected to calculate an osRfC of 0.01 mg/m3 for TTP. A UFC of 10 was applied to address UFH.*

NIEHS Comments:

Tier 1

- Based on the summary of the available data, it is reasonable to conclude that female reproductive toxicity based on TTP (Taskinen et al. 1999) and spontaneous abortion (John et al. 1994, Taskinen 1994, 1999) is supported by epidemiological evidence. No additional comments to suggest.
- It would likely be appropriate to state that by inference, one likely contributor to the observed toxicity has been extensively characterized. It has been clearly demonstrated that formaldehyde is metabolized to formic acid; formic acid has effects on embryo viability and developmental outcomes. *Teratology*. 1995 Apr;51(4):243-51. doi: 10.1002/tera.1420510409; *Cell Biol Toxicol.* 2004 May;20(3):133-45. doi: 10.1023/b:cbto.0000029466.08607.86. It would likely be appropriate to state that by inference, one likely contributor to the observed toxicity has been extensively characterized. It has been clearly demonstrated that formaldehyde is metabolized to formic acid; formic acid has effects on embryo viability and developmental outcomes. *Teratology*. 1995 Apr;51(4):243-51. doi: 10.1002/tera.1420510409; *Cell Biol Toxicol.* 2004 May;20(3):133-45. doi: 10.1023/b:cbto.0000029466.08607.86
- This appears to be a logical selection of POD and uncertainty factor. No comments to suggest.

b. *Male reproductive toxicity*
    - *The assessment concludes that the evidence indicates that inhalation of formaldehyde likely causes reproductive toxicity in men, given the appropriate exposure circumstances. The conclusion for male reproductive toxicity is supported primarily by coherent evidence of several alterations to the male reproductive system in animals exposed to very high levels of formaldehyde (>6 mg/m3), with some corroborative changes in an occupational epidemiological study; although no MOA is available, some relevant mechanistic changes have been observed in well-conducted studies of the male reproductive organs of exposed rodents.*
    - *A POD from Özen et al. (2002), a rat study, was ultimately selected to calculate an osRfC of 0.01 mg/m3 for testis weight. A UFC of 3,000 was applied to address UFH, LOAEL (UFL), UFS, and UFA.*

NIEHS Comments:

Tier 1

- The data presented in J. Trace Etem. Med. Biol. VoL. 16, pp. 119-122 (2002) is of low quality, N=7. This is too small of a number to generate data that one could be confident in. These weights were also presented as a relative weight (which weight-terminal? Body weight gain?). The apparent absence of this clarity decreases the confidence of the study findings. Given that T is secreted in a pulsatile manner, an N of at least 10 is typically utilized. Samples must also be collected in a short duration. In contrast, the study by Vosougi et al. utilized 6/group. Decreases in T were observed in both studies, but the weight (absolute presented) did not show a change. The Vosougi study has a similar N. Given the photomicrographs of <u>properly fixed</u> testes, the histopathology demonstrates an adverse effect at 10 ppm, in addition to effects on sperm parameters, collectively supports the *High* confidence classification.

- Table 1-35. Confirm that the 2005 Ozen study examined histopathology because data on that endpoint were not found in the assessment. The information provided may instead refer to the 2008 publication by the same Author. Nonetheless, fixation in formalin greatly decreases the confidence in the resultant histopathology.
- The conclusion that formaldehyde likely causes reproductive toxicity in men, as supported/corroborated by an occupational epidemiology study (Wang et al. 2012), is a reasonable conclusion. No additional comments to suggest.

Tier 2

- P 1-387 line 27. Consider reviewing the manuscript *Food Chem Toxicol.* 1989 Aug;27(8):545-8. doi: 10.1016/0278-6915(89)90051-3. The authors demonstrated decreases in fetal weight at 20 and 40ppm (0,5,20, 40 ppm assessed). This is a quite robust guideline-like study.
- Tables (in general). It would be of benefit to the reader if all columns were displayed consistently (e.g., numbers in different rows in a column should appear clearly and centered, and collective study summary data not inappropriately 'break' with a page break and no headers, shading, or lines by study (e.g., Table 1-26). The use of "%" is also inconsistent (e.g., Table 1-46 Sheveleva (71) and Usanmaz (2002)).
- P 1-367-368. The medium confidence study by Taskinen et al. (1999) is robustly discussed, but no mention was given to the more recent, low confidence Danish cohort study by Zhu et al. (2005)., Suggest adding a brief summary of the strengths/weaknesses and the null results. This would be consistent with the following section on spontaneous abortion, which summarizes a low confidence study.
- P 1-370, line 9-12. It is unclear whether the IRIS method reports both that the less precise exposure assessment method increases the likelihood of (likely non-differential) exposure misclassification (information bias) and the poorer study sensitivity.
- In the descriptions of epidemiological studies, there is no indication (except in the Tables) of the study design. In discussing concepts like recall bias, which is mostly present in case-control studies, it would be helpful to state the study designs.
- P 1-411, line 34. Provide a space between "group" and "because".
- P 1-412, line 5. Provide a reference for the third study (Taskinen et al.).

c. Nervous system toxicity

- *While many studies reporting evidence of potential neurotoxic effects were available—including developmental neurotoxicity, multiple manifestations of behavioral toxicity, and an increased incidence of, or mortality from, the motor neuron disease amyotrophic lateral sclerosis—due to limitations identified in the database (e.g., poor methodology, lack of consistency), it was ultimately determined that the evidence suggests, but is not sufficient to infer, that formaldehyde inhalation might pose a human health hazard. The evidence integration narrative emphasizes that additional study is warranted.*
- *The available data on potential nervous system effects were considered insufficient for developing quantitative toxicity estimates.*

NIEHS Comments: The determination that due to limitations identified in the database, the evidence suggests, but is not sufficient to infer that formaldehyde inhalation might pose a human health hazard is appropriate.

Human evidence

- **Formaldehyde exposure and neurobehavior summary:** The designation of low confidence to three observational and two controlled exposure studies, and medium confidence to one controlled exposure study is appropriate. Results suggesting that formaldehyde exposure might be associated with deficits in performance in neurobehavioral tests related to memory, coordination and motor control were inconsistent and potentially confounded by co-exposure to other neurotoxicants (e.g., methanol containing formalin), irritation and differences in population characteristics (e.g., age, education). Some studies were also missing important details (e.g., data not shown). Additional studies are needed that directly address formaldehyde exposure and account for potential confounding or mediating factors.

- **Formaldehyde exposure and death from ALS summary:** The designations of high confidence to one prospective cohort study and medium confidence to the other six observational studies (2 prospective cohort and 4 case-control) is appropriate. Results suggesting that exposure to formaldehyde and death from ALS were inconsistent, which may be attributed to the small numbers of exposed cases and uncertainty in individual exposure assignments. These are areas that should be addressed in future work.

Animal evidence

- Recognition that the identified limitations in the animal studies (i.e., poor methodology, lack of consistency, compound specificity) warrant additional studies, especially chronic and developmental exposures is appropriate.
- **Nervous System Effects in Animal Studies**: The designation of low confidence to 15 exposure studies, and medium confidence 3 exposure studies is appropriate. Results suggesting that formaldehyde exposure might be associated with deficits in performance in neurobehavioral tests related to developmental pathology, sensitization, learning/memory, and motor activity were inconsistent and potentially confounded by co-exposure to other neurotoxicants (e.g., methanol containing formalin), sample size/litter analysis, irritation, and differences in timing of behavioral testing. Some studies were also missing important details (e.g., data not shown) or not specifically designed to access nervous system effects so excluded (i.e., one sub chronic and three chronic). Additional well-controlled animal studies are needed that directly address formaldehyde exposure and designed to mimic the human condition.

Tier 2

- Possible error in Table 1-44. Only two of the six studies have figures showing the confidence ratings.
- p. 1-335 line 4. Add reference to ADME for following statement, "As appreciable amounts of formaldehyde are not expected to reach the systemic circulation or CNS to elicit direct effects, any potential mechanisms would need to be indirect."
- Tables are not referenced correctly in appendix.
    - Examples on p. A-583.
    - Line 7, "The search strings used in specific databases are shown in Table A-82." This is table A-83.
    - Line 26, "Inclusion and exclusion criteria used in the screening steps are described in Table A-83." This is table A-84.
    - Line 27, "The search and screening strategy, including exclusion categories applied and the number of articles excluded within each exclusion category, is summarized in Figure A-37." This figure is A-34.

6. Cancer

*The assessment concludes that formaldehyde is **Carcinogenic to Humans by the Inhalation Route of Exposure**. Please comment on whether the judgments in (a) to (e), below, are clearly described and scientifically justified. Note that the judgments in (a) and (b) outline the primary support for this conclusion across two lines of evidence, each of which would independently substantiate the carcinogenicity conclusion.*

a. *The **evidence demonstrates** that formaldehyde inhalation causes nasopharyngeal cancer (NPC) in humans, based on observations of increased risk of NPC in groups exposed to occupational formaldehyde levels and nasal cancers in animals, with strong, reliable, and consistent mechanistic evidence in both animals and humans (i.e., robust evidence for both the human and animal evidence, and strong mechanistic support for the human relevance of nasal cancers observed in animals).*

NIEHS Comments: The collective evidence demonstrates that exposure to formaldehyde causes NPC based on the rationale provided above and that "the available epidemiological studies provide robust evidence of an association consistent with causation between formaldehyde exposure and increased risk of nasopharyngeal cancer." EPA has

done an outstanding job in systematically evaluating the evidence and applying Hill guidelines to the body of evidence.

Tier 1

- Study selection
  - Exclude studies of occupations that do not provide formaldehyde-specific risk estimates and potentially involve exposure to other carcinogens. These studies include Yu et al. 2004 (link with formaldehyde is speculative), Vaughan 1989 (carpenters), and possibly Malker et al. 1990 (fiberplant is potentially relevant, but the risk estimate is for industry rather than occupation). (Studies of health professionals, such as embalmers, are considered relevant because they are exposed to high levels of formaldehyde and lower potential for confounding from co-exposures.)
  - Replace Hayes et al. 1990 with Hauptman et al. 2009 (a more informative study that supersedes Hayes). Although Hayes reported an excess of NPC (based on four cases), Hauptmann found that two of the four NPC cases were not among embalmers: OR for "ever embalming" = 0.1, 95% CI = 0.01 to 1.2.
- Study evaluation
  - Recommend upgrading Vaughan et al. 2000 from medium confidence to high confidence. NRC (2014) considered this study to have a high-discrimination quantitative exposure assessment.
- Evidence evaluation/integration
  - Strength of the association. Use caution in reporting the number of positive studies with a specific magnitude because many risk estimates are imprecise (e.g., small numbers of exposed cases). Thus, while there is high confidence that there is a positive association between formaldehyde exposure and NPC risk, there is lower confidence in the exact magnitude of the risk estimate.
  - In the NCI cohort of formaldehyde workers, five of the 10 cases occurred in plant 1. Because this issue has been a criticism of the literature and the NCI is one of the most informative studies, please discuss the author's influence analysis to evaluate the consistency of findings across plants.
- Summary figure (1-20)
  - Stratify the risk estimates by study confidence (similar to the myeloid leukemia summary plot) rather than cell type; study quality is more informative for exploring potential heterogeneity and providing transparency for the conclusions.
  - Consider removing the non-informative cohort studies with zero cases from the summary figure.
  - Ideally, only one estimate (using a similar metric across studies, such as highest exposure) should be plotted for each independent study population (unless there are multiple subpopulations, such as men/women). (Exposure-response plots and related p values for the relevant studies could be plotted as a matrix below the summary forest plot or as a separate figure.)

Tier 2

- Study inclusion
  - For completeness, briefly mention the results of Doll and Teta (0 cases of NPC in 111 formaldehyde exposed workers, see text in publication).
- Study evaluation:
  - Consider upgrading Hildesheim et al. 2001 from low confidence to medium confidence. Table 1-32 (EPA report) states that the analysis did not control for smoking; however, Hildesheim states (methods) that "Additional adjustment for family history of NPC, cigarette smoking, ...did not materially affect risk estimates (data not shown)."
  - Create a new study confidence category, low confidence, because of inadequate statistical power to detect very rare cancers. This new category would increase transparency for plotting or removing the uninformative cohort studies (e.g., zero cases) from the summary figure. It

would also help distinguish low confidence studies resulting from bias from methodologically sound smaller studies. Note that cohort studies that did not report on NPC most likely did not observe any cases.

- Summary figure (1-20)
  - The attempt to present the data in summary plots is appreciated; however, the figures are hard to read due to the small font size and vertical text.
- Reporting and other comments
  - Suggest moving the appendix exposure assessment tables to the main document.
  - The discussion of cancer incidence and tumor subtype across geographical regions could be integrated better (e.g., it is scattered several places in the discussion).
  - Page 1-198. Recommend deleting the sentence that EPV is not a sufficient cause of cancer as that may be true of almost all carcinogens, including smoking (see Rothman and Greenland, 2005)

b. *The **evidence demonstrates** that formaldehyde inhalation causes an increased risk of myeloid leukemia in humans, based on robust human evidence from observations of increased risk in groups exposed to occupational formaldehyde levels. This judgment is supported by other studies of human occupational exposure that provide strong and coherent mechanistic evidence identifying clear associations with additional endpoints relevant to lymphohematopoietic (LHP) cancers, including an increased prevalence of multiple markers of mutagenicity and other genotoxicity in peripheral blood cells of exposed workers, other perturbations to immune cell populations in blood (primarily from human studies), and evidence of other systemic effects (i.e., developmental or reproductive toxicity). Generally, evidence supporting the development of LHP cancers after formaldehyde inhalation has not been observed in experimental animals (i.e., rodents), including a well-conducted, chronic cancer bioassay in two species, a similar lack of increased leukemias in a second rat bioassay, and multiple mechanistic evaluations of relevant biological changes such as genotoxicity in systemic tissues of exposed rodents (resulting in a judgment that the animal evidence is indeterminate).*

NIEHS Comments: The collective evidence demonstrates that exposure to formaldehyde causes myeloid leukemia based on the rationale provided above and also that "the available epidemiological studies provide robust evidence of an association consistent with causation between formaldehyde exposure and increased risk of myeloid cancer." EPA has done an outstanding job in systematically evaluating the evidence and applying Hill guidelines to the body of evidence.

Tier 1

- Study selection
  - Exclude Pira et al. 2014 because it is not specific to formaldehyde and workers were exposed to carcinogens such as styrene.
- Evaluation/evidence integration
  - Counting studies should be based on independent populations and not publications because several reports analyze the same or overlapping study populations. The EPA document mentions the overlap of the embalmer reports (e.g., the case-control study by Hauptmann et. al., 2009 is nested in the combined population from the three cohort studies by Hayes, and Walrath and Frauemi 1984, 1983); however, the discussion often counts these reports as four studies. It may not be necessary to discuss the findings from the three cohort studies in the text unless they provide additional information (such as specific types of myeloid leukemia).
  - The statement that all 13 informative studies found an elevated risk of formaldehyde above the null is misleading because it double counts studies and includes a non-formaldehyde-specific and a null study (Saberi Hosnijeh). Although the risk estimates for some of the other studies were technically above one, the findings are unclear because of very imprecise estimates resulting from the small number of exposed cases or inconsistencies (e.g., Blair 2001, Ott et al. 1989).

- A clearer discussion of latency and the time-period analysis (e.g., extending follow-up by yearly increments) in the NCI study (Beane Freeman et al. 2013, Figure 1) is needed. The EPA report notes that this analysis shows consistent findings of a strong relationship between myeloid leukemia and formaldehyde exposure over the entire follow-up period. However, EPA should also discuss that the overall association between peak formaldehyde exposure and myeloid leukemia has decreased with an additional 10-year follow-up (e.g., 2004 update compared to the 1994 update). This risk pattern is consistent with a shorter induction period for myeloid leukemia and a shorter latency period has been observed for other leukemogenic chemicals. (Note that exposure in the cohort probably ceased in 1980). Based on these findings, the 1994 update may be more informative than the 2004 update and should be plotted in the summary figure (instead of the 2004 findings). The NIOSH cohort findings were also somewhat weaker in the latest follow-up. EPA should consider adding the results from the earlier report (Pinkerton 2004) of this cohort to the Meyers entry in Table 1-60 and discussing the results in the text (as part of the discussion of Meyers) if needed.

- Summary figures (1-37, 1-38)
  - Consider deleting overlapping studies (Hayes, Walrath and Fraumeni 1983, 1984, which overlap with Hauptmann) and Checkoway, which overlaps with Beane freeman 2009.
  - Consider deleting non-formaldehyde specific study (Pira 2014).
  - Figure 1-37: Stratify risk estimates by study confidence (similar to Figure 1-38) instead of population-level exposure vs individual assessment (as only one study in population-level assessment).
  - Add a footnote with *p* values for trends, as the strength of the evidence shows statistically significant positive trends in addition to pair-wise comparisons.
  - Ideally, only one estimate (using a similar metric across studies, such as highest exposure) should be plotted for each independent study population (unless there are multiple subpopulations, such as men/women). (Exposure-response plots and p values for the relevant studies could be plotted as a matrix below the summary forest plot or as separate figure.)

Tier 2

- Page 1-425, lines 11 to 18. State that formaldehyde was associated with AML an CML in the EPIC study (Saberi Hosnijeh et al., 2013); according to Table 3 of the study, the OR was 1.02 for AML and 0.92 for CML.
- Figure 1-39 is not very informative (correlation of AML with myeloid leukemia).
- Consider reporting the results of the meta-analysis of high exposure to formaldehyde and myeloid leukemia (Schwik et al. 2010).
- Summary figure is hard to read because of small font size and vertical text.

c. *This carcinogenicity conclusion is corroborated by several other lines of evidence for which the **evidence indicates** that formaldehyde inhalation likely causes that cancer type in humans, namely sinonasal cancer, oropharyngeal/hypopharyngeal cancer, and multiple myeloma.*
d. *Formaldehyde is genotoxic in several test systems and operates, at least in part, through a mutagenic MOA. Specifically, a mutagenic MOA was identified in association with the development of nasal (including nasopharyngeal and sinonasal) cancers, while a mutagenic MOA was not identified for other cancer types. The mechanistic evidence was sufficient to conclude that both mutations and cellular proliferation play a role in nasal carcinogenesis.*
e. *The exact mechanism(s) leading to cancer formation outside of the respiratory tract are unknown.*

NIEHS Comments: EPA's conclusion regarding genotoxicity and other mechanistic data for respiratory and non-respiratory cancers (d and e) is appropriate. The systematic evaluation and assessment of the evidence are scientifically sound and clearly described. The conclusions(c) for the other cancer outcomes are less clear. The collective evidence – cancer and mechanistic studies in humans, experimental animals, and cells – is more robust for SNC than for the other cancers (oropharyngeal/hypopharyngeal cancer, and multiple myeloma). The conclusion that

the evidence from human studies is indeterminate or slight for laryngeal cancer, lymphatic leukemia, and Hodgkin lymphoma based on inconsistent findings across studies is appropriate.

Sinonasal Cancer (SNC)

Tier 1

- Rationale for conclusions
    - Given that EPA has concluded that (1) formaldehyde is genotoxic in humans (based on studies using buccal, blood, and nasal tissues from exposed humans) and in experimental systems, (2) a genotoxic mode of action plays a role in the development of nasal tumors (SNC) and NPC, and (3) there is moderate confidence for a causal association between formaldehyde exposure and SNC risk from human studies, EPA should discuss why the collective evidence from human cancer and mechanistic studies does not support robust evidence for a causal relationship between SNC and formaldehyde exposure
    - EPA should provide a more transparent rationale for why the human evidence is moderate instead of robust. EPA's rationale mainly focuses on the evidence supporting an association – consistency, strength of the association, and evidence of an exposure-response relationship. Moreover, EPA concludes that there is reasonable confidence that alternative explanations have been addressed, including chance, bias, and confounding within individual studies or across studies; this is usually the criteria for demonstrating causality. (Residual confounding from confounding could be a reasonable rationale for moderate evidence). The text states that many analyses lack precision due to the rarity of the tumors. However, statistically significant findings were found in the pooled analysis of 12 studies in different countries and occupations, which provides evidence of consistency similar to that found from multiple studies. EPA also noted that there is some uncertainty due to the lack of data on the temporal relationship allowing time for cancer induction, latency, and mortality. However, this seems to be a weak argument. Cancer latency may be difficult to assess especially given the fact that wood dust may be an effect modifier for formaldehyde-induced SNC.
- Study selection
    - Consider excluding formaldehyde nonspecific studies: Teschke et al. 1997, Jakobsson et al.
- Study confidence evaluation
    - Consider upgrading Luce 2002 from medium confidence to high confidence. Note that NRC (2014) classified this study as having a high-discrimination quantitative exposure assessment. Although the exposure assessment was based on job title/industry, industrial hygienists used industrial hygiene data to develop the JEM and create semi-quantitative exposure groups.
    - Besides including prevalent cases, Pesch's study has several other methodological issues, e.g., the cases had occupational diseases, whereas the controls were involved in accidents and all subjects (cases/ and controls) were woodworkers. These issues may limit the ability of the study to detect a true effect.
- Evaluation/evidence integration
    - Consider discussing the results from the individual studies in Luce et al. 2002 pooled analysis (e.g., Luce 1993, Hayes 1986) as part of the discussion of the Luce study (without double counting). For example, Luce 1993 evaluated average, cumulative, and duration, and Hayes found similar findings using two different industrial hygienists' assessments. These findings help increase the confidence of the body of literature.
    - Consider discussing the potential effect modification by wood dust.
- Summary figure (1-21)
    - Consider stratifying the summary figure by study confidence rather than by cell type. Study quality is informative for exploring heterogeneity and providing transparency for the conclusions.
    - Consider removing the findings from the non-informative cohort studies with 0 cases from the figure.

- Ideally, only one estimate (using a similar metric across studies, such as highest exposure) should be plotted for each independent study population (unless there are multiple subpopulations, such as men/women). Exposure-response plots for the relevant studies could be plotted as a matrix below the summary forest plot or as a separate figure.

Tier 2

- Study inclusion
  - For completeness, mention the findings from Doll and Teta (0 cases of NPC in 111 formaldehyde exposed workers, see text).
- Study evaluation
  - As per the NPC recommendation, consider creating a new study confidence category, low confidence, because of inadequate statistical power to detect very rare cancers. This new category would increase transparency for plotting or removing the uninformative cohort studies (e.g., zero cases) from the summary figure. It would also help distinguish low confidence studies resulting from bias from methodologically sound smaller studies. Note that more cohort studies were reported for SNC than NPC and thus, there may be a similar number of cohort studies with zero cases of NPC.
- Summary figure (1-21)
  - The figure is hard to read because of the small font size and vertical text.

Oropharyngeal/hypopharyngeal cancer

Tier 1

- Study selection
  - Consider excluding studies that are not specific for formaldehyde exposure, e.g., Vaugh et al. 1989.
- Evaluation/integration
  - As per previous comments, study counts should be for unique populations and not reports.
  - Recommend not calculating SMR from the Coggon study because the internal analysis was provided, which is more informative than the external analysis. Potential confounding from smoking and alcohol consumption are greater concerns in an external than an internal analysis because workers may smoke more than the general public.
  - EPA should discuss potential methodological issues in Marsh's (2002/2007) re-analysis of the Wallingford plant included in the NCI cohort of formaldehyde workers. EPA notes that confounding from co-exposures is not a concern because Beane Freeman did not find evidence of potential confounding from multiple co-exposures (asbestos is a risk factor for pharyngeal cancer) in the NCI study. However, it is unclear whether the findings from internal analysis using workers as the referent group are relevant to the external analysis reported in the Marsh report. Beane Freeman also stated that smoking rates were high in the NCI cohort but not a potential confounder in internal analyses because rates did not vary by exposure level; however, smoking could be a confounder in external analyses. The evaluation of the magnitude of the risk in the different formaldehyde exposure categories is complicated by the elevated SMR in the non-exposed workers. Evidence of an exposure-response relationship is less clear in the Marsh study. Duration of response was based on the small number of exposed cases and was not apparent in analyses of higher exposure levels. Another potential concern is the 10-fold differences in predicted exposure levels assessment in the Marsh vs. NCI reports.
  - EPA states (1-242) that smoking and alcohol consumption are unlikely to be generally related to occupational and residential formaldehyde exposures and are, therefore, unlikely to be across-the-board confounders. Smoking and alcohol consumption may be less of a concern for internal analyses of workers but could be a concern for external analyses.
  - EPA recommendation of moderate evidence is not clear. It is mainly driven by the positive findings in the LaForest's study. The Marsh study is difficult to interpret and findings from other studies are null or weak.

- Summary figure (1-22)
    - Consider stratifying the summary figure by study confidence rather than by cell type. Study quality is informative for exploring heterogeneity and providing transparency for the conclusions.
    - Graphing the SMR Marsh studies may distort actual risk in the studies as there is excess risk in the unexposed group.
    - Ideally, only one estimate (using a similar metric across studies, such as highest exposure) should be plotted for each independent study population (unless there are multiple subpopulations, such as men/women). Exposure-response plots for the relevant studies could be plotted as a matrix below the summary forest plot or a separate figure.

Tier 2

- Summary figure (1-22)
    - This figure is hard to read because of the small font size and vertical text.

Multiple myeloma

Tier 1

- Study selection
    - Recommend excluding studies that are not specific for formaldehyde exposure, e.g., Pira, Band et al. 1997, Dell and Teta (multiple myeloma not reported for 111 formaldehyde exposed workers).
- Evaluation/evidence integration
    - Use caution in counting studies as individual/unique populations as per previous comments. Hayes (could be deleted from the summary of the results) overlaps with Hauptmann and Stellman 1998 and with Boffetta 1989 (it is unclear if the four formaldehyde cases are the same in each study). Pottern (Danish women) and Heinemann (Danish men) are also from the same population base. Counting studies is also not informative since many studies have few exposed cases.
    - Page 1-462 lines 1 to 7, the text is unclear: "Seven of the 14 informative studies reported increased risk of death from multiple myeloma associated with exposure to formaldehyde (Hauptmann et al., 2009; Band et al., 1997; Dell and Teta, 3 1995; Heineman et al., 1992; Pottern et al., 1992; Boffetta et al., 1989; Edling et al., 1987b). Four studies reported mixed or null results (Coggon et al., 2014; Meyers et al., 2013; Beane Freeman et al., 2009; Ott et al., 1989), and three studies reported a decreased risk of death from multiple myeloma associated with exposure to formaldehyde (Pira et al., 2014; Stellman et al., 1998; Band et al., 1997)." Of the seven studies reported showing increased risk, two are not specific for formaldehyde (Dell and Teta and Band), three have less than four cases (Pottern, Boffetta, Edling), and the other two have non-significant risk estimates of 1.1 and 1.4. In contrast, the strongest evidence for an association between formaldehyde and multiple myeloma is from the NCI study (Beane Freeman), which is categorized as unclear in this sentence, albeit the evidence is better described later in the paragraph.
    - Overall, the strength of the association across studies is somewhat weak; the only significant finding (from any analyses) is for the highest peak exposure category in the NCI cohort; test for trend = 0.08 (exposed only). Elevated risk estimates from the other studies are imprecise and usually less than 1.5 or based on small numbers. Although statistical significance is just one factor to consider, more consistent results from more studies may be needed to reach a moderate level of evidence.
- Summary figure (1-401, should this be 1-41?)
    - Consider stratifying the summary figure by study confidence, which is informative for exploring heterogeneity and providing transparency for the conclusions. The stratification of the 3 studies by presumed peak exposure is not informative as the studies may not be specific for

formaldehyde, have small numbers, or overlap with Hauptmann (could replace Hayes with Hauptmann)
    - Consider removing the findings from the non-formaldehyde specific studies and from the overlapping studies (e.g., Hayes).
    - Ideally, only one estimate (using a similar metric across studies, such as highest exposure) should be plotted for each independent study population (unless there are multiple subpopulations, such as men/women). Exposure-response plots for the relevant studies could be plotted as a matrix below the summary forest plot or as a separate figure.

Tier 2

- Summary figure (1-401, should this be 1-41?)
    - This figure is hard to read because of the small font size and vertical text.

Mechanistic studies

Tier 2

- Consider moving Table A-27 (Genotoxicity summary) from the appendix to the main document.

Tier 3 (all cancer outcomes)

- Potential environmental health disparities remain a research gap.

Note: NIEHS did not provide comments to questions 5 (Rfc for non-cancer) and 7 (inhalation unit risk for cancer).