



Final Report for Letter Peer Review of Ethylene Oxide Carcinogenic Dose-Response Assessment Development Support Document

Contract Number 582-19-91305

Work Order Number: 582-19-91305-01 (PCR
11740)

Prepared by:
Risk Science Center
Department of Environmental and Public Health Sciences
University of Cincinnati

Prepared for:
Sabine Lange, Ph.D.
Texas Commission on Environmental Quality

Contact:
Jacqueline Patterson (Jacqueline.patterson@uc.edu)

April 30, 2020

Note

This report was prepared by scientists from the University of Cincinnati Risk Science Center (RSC). The peer reviewers served as individuals, representing their own personal scientific opinions. They did not represent their companies, agencies, funding organizations, or other entities with which they are associated. Their opinions should not be construed to represent the opinions of their employers or those with whom they are affiliated.

Table of Contents

| | |
|--|----|
| TABLE OF CONTENTS | 3 |
| INTRODUCTION AND BACKGROUND..... | 5 |
| Charge to Peer Reviewers | 6 |
| Clarifying Questions on Draft Responses to Charge Questions..... | 8 |
| EXPERT PEER REVIEWER RESPONSES TO CHARGE QUESTIONS | 10 |
| Charge Question 1 | 10 |
| Charge Question 2 | 12 |
| Charge Question 3 | 16 |
| Charge Question 4 | 22 |
| Charge Question 5 | 28 |
| Charge Question 6 | 30 |
| Charge Question 7 | 33 |
| Charge Question 8 | 37 |
| Charge Question 9 | 40 |
| Charge Question 10 | 44 |
| Charge Question 11 | 47 |
| Charge Question 12 | 50 |
| Charge Question 13 | 51 |
| REFERENCES | 54 |
| APPENDIX A – PEER REVIEW ORGANIZATION..... | 57 |
| APPENDIX B - CHARGE QUESTIONS AND INSTRUCTIONS..... | 61 |

APPENDIX C –BIOSKETCHES AND CONFLICT OF INTEREST SCREENING..... 66

APPENDIX D –REFERENCES IDENTIFIED BY EXPERTS..... 73

Introduction and Background

This report provides the final peer review comments from six independent experts on an ethylene oxide (EtO) carcinogenicity assessment prepared by the Texas Commission on Environmental Quality (TCEQ). The University of Cincinnati Risk Science Center (RSC) organized this independent letter peer review under a contract to TCEQ.

The Toxicology, Risk Assessment, and Research Division of TCEQ conducted a systematic review and dose-response assessment of ethylene oxide (EtO) carcinogenicity and prepared a Development Support Document (DSD) entitled “Ethylene Oxide Carcinogenic Dose-Response Assessment.” TCEQ toxicity assessments and values are used in the evaluation of air permit applications and ambient air data and are developed following *RG-442 TCEQ Guidelines to Develop Toxicity Factors* (TCEQ, 2015). A draft DSD was issued by TCEQ in June of 2019 for public comment. TCEQ received numerous public comments from diverse groups and individuals. It reviewed and considered the public comments and revised the DSD. It also prepared a response to public comments document. These two documents are the subject of this peer review.

The purpose of this peer review was to provide TCEQ with expert opinions on the TCEQ inhalation cancer unit risk factor (URF) for EtO, in order to determine if it is scientifically adequate and appropriate for estimating cancer risk at ambient (low-level) concentrations. Experts were also asked to comment on the clarity and transparency of the two documents.

The RSC organized this peer review and was responsible for managing all aspects of the process, including selection of the reviewers, evaluation of potential conflicts of interest, development of the charge questions, distribution of the assessment document, collection and review of each expert’s written comments, and collation of all comments into this report. RSC used best practices for peer review that reflect guidance from the Office of Management and Budget, the National Academy of Sciences, the U.S. Environmental Protection Agency (US EPA) and others. A description of the process used to organize and conduct this peer review is found in Appendix A. The instructions for peer reviewers, charge questions, and a list of the documents included in the review package is found in Appendix B.

RSC independently selected a group of six experts to provide a diversity and balance of relevant expertise and backgrounds. The group included at least two individuals who are expert in each of these key areas: environmental epidemiology; cancer dose response modeling; and, cancer risk assessment/toxicology/mode of action. The selected expert peer reviewers for this review are listed below. Their affiliations are provided for identification purposes only. Biographical sketches and information with regard to conflict of interest screening are found in Appendix C.

- Bruce Allen, M.S., Biomathematician, independent consultant
- Kenny Crump, Ph.D., Biomathematician, private consultant, retired
- Harvey Checkoway, Ph.D., Professor of Epidemiology, Department of Family Medicine and Public Health and Department of Neurosciences, University of California, San Diego
- Bette Meek, Ph.D., Associate Director of Chemical Risk Assessment, McLaughlin Centre for Risk Science, University of Ottawa

- David Savitz, Ph.D., Professor of Epidemiology and Interim Chair, Department of Epidemiology in the School of Public Health, and Obstetrics and Gynecology and Pediatrics in the Alpert Medical School, Brown University
- Rita Schoeny, Ph.D., Consultant in risk assessment and science policy, retired from US EPA

RSC collated the expert reviewers' final comments into this report, which is organized by charge question so that one can see the reviewers' comments and opinions on each charge question together. The experts were randomly assigned a number to identify their responses, allowing the specific comments to remain anonymous to TCEQ and the readers of this report. This was done to make the experts feel more comfortable in expressing their candid opinions. This final report constitutes the record of this peer review. A copy of the final report will be made available to the public on the RSC web site (<https://med.uc.edu/eh/centers/rsc/peer-review>).

Charge to Peer Reviewers

Background and Timeline

The following two paragraphs were provided by TCEQ to provide background for the experts:

Ethylene oxide (EtO) is used as a chemical intermediate in the manufacture of ethylene glycol (antifreeze), polyester, detergents, polyurethane foam, solvents, medicine, adhesives, and other products. Relatively small amounts of EtO are used in sterilization of surgical equipment and plastic, as a fumigant, and as a sterilant for food (spices) and cosmetics. In 2018, EtO was being produced in the US at 15 facilities in 11 locations by nine companies. In the US, EtO is primarily produced in Texas and Louisiana. IARC has designated EtO as a group 1 human carcinogen (IARC, 2012). Between October 1, 2018 and March 31, 2019, the EPA conducted air monitoring for EtO in various locations in the United States, and found that the levels of EtO concentrations that are considered to be "urban background" are in the range of 0.1 – 0.2 ppb (<https://www.epa.gov/hazardous-air-pollutants-ethylene-oxide/ethylene-oxide-data-summary-national-air-toxics-trends>). In regard to longer-term levels around EtO-emitting facilities, as an example, the mean and 95th percentile modeled 5-year concentrations for one sterilizer facility were ≈0.17 and 1.6 ppb, respectively ([https://www.atsdr.cdc.gov/HAC/pha/sterigenic/Sterigenics International Inc-508.pdf](https://www.atsdr.cdc.gov/HAC/pha/sterigenic/Sterigenics_International_Inc-508.pdf)).

In early 2017, as part of a standard yearly review of newly-derived toxicity factors, the TCEQ Toxicology Division reviewed the EPA's cancer-based toxicity factor derivation for EtO (finalized in 2016) to determine if we would provisionally adopt the EPA's number for use in deriving protective concentration levels (PCLs) for the Texas Risk Reduction Program (TRRP). In March 2017 the division decided that, instead of adopting the EPA's EtO toxicity factor, we would derive an interim EtO toxicity factor for the agency's use in the remediation program with a plan to conduct a complete evaluation of EtO inhalation carcinogenicity for use in both air permitting and remediation. The TCEQ decided to complete this thorough evaluation because EtO is emitted in Texas and has been determined to be a carcinogen. In August of 2017, the TCEQ announced a 90-day public information request for scientific information about EtO that may be of use in the TCEQ's review. The TCEQ then completed a systematic review and dose-response assessment of EtO carcinogenicity and released the draft Development Support Document (DSD) on June 28, 2019 for public comment. The public comment period ended in late September 2019 and the TCEQ undertook a review and response to the public comments and scientifically

justified revisions to the draft DSD. The revised DSD and response to comments are the documents that are under review in this peer review.

The following is a list of the charge questions for this peer review.

Charge Questions

1. The TCEQ conducted a systematic review of the literature relevant to the derivation of an inhalation unit risk factor for ethylene oxide (EtO) (see DSD Appendix 1). Are you aware of any additional literature or studies that should be considered and if so, how might they impact the assessment?
2. The TCEQ adopts the EPA conclusion that the weight of the evidence supports a direct-acting mutagenic mode of action (MOA) for EtO carcinogenicity (DSD Sections 3.3 and 3.3.1). Section 3.3.1 of the DSD presents summary information from the EPA (EPA Section 3.4.3) relevant to the MOA determination. Do you agree with the MOA determination? Please explain.
3. The TCEQ adopts EPA's MOA analysis (DSD Section 3.3.1) and considers MOA as information relevant to the likely or expected shape of the dose-response (DSD Sections 3.4.1 and 3.4.1.1) as specified by the TCEQ guidelines for developing toxicity factors (TCEQ, 2015). What is your opinion on whether and how the MOA should inform the likely or expected shape of the dose-response curve, overall and in the low-dose range (e.g., at environmentally-relevant concentrations); and whether and how the MOA should inform the choice of dose-response model for estimating human carcinogenicity risk? Please comment on TCEQ's reasoning on the implications of the MOA for the shape of the dose-response and its relative importance amongst their other model choice considerations (summarized in DSD Section 3.4.1.4.2). Are the TCEQ conclusions concerning implications of the MOA scientifically defensible?
4. The TCEQ conducted an evaluation of EtO's carcinogenic classification (DSD Section 3.3.2), and also evaluated breast cancer risk in humans as a potential cancer endpoint (DSD Appendix 6; Response to Dr. Kyle Steenland, Comment 1 in Response to Public Comments Document). What is your characterization of the overall weight of the evidence for or against EtO increasing the risk of breast cancer in humans at occupational concentrations (past or present) and at environmentally-relevant concentrations?
5. While it is in the interest of public health to protect against cancer *incidence*, available epidemiological studies often only provide cancer *mortality* data for dose-response modeling. What is your opinion on the accuracy of using a dose-response model based on cancer mortality data (e.g., lymphoid cancer mortality) to predict cancer incidence (e.g., lymphoid cancer incidence)?
6. The TCEQ's DSD discusses a problem with key USEPA AIC and p-value calculations used as criteria in determining model fit, and the TCEQ recalculated these values (DSD Section 3.4.1.3 and Appendix 4). Please explain what you think the appropriate approach should be for accounting for the number of estimated parameters in the modeling and the associated calculation of the AIC and p-values. Given that appropriate AIC and p-values are available for models fit to *individual* data, what role should visual fit to *categorical* estimates play in model selection (Response to University of California at San Francisco, Comment 6 in Response to Public Comments Document)?

7. Please comment on the biological and mechanistic support for and against use of an overall supralinear model to estimate risk of lymphoid cancer from exposure to EtO at occupational levels and at environmentally-relevant concentrations.
8. As summarized in DSD Section 3.4.1.4.2, the TCEQ used MOA, model predictiveness reality checks (both for the NIOSH cohort and the general population), biological plausibility, and statistical model fit criteria for model selection. Have these considerations been clearly described and are they scientifically appropriate given the available data?
9. In DSD Sections 3.4.1.4, 3.4.1.5, and 3.4.1.6, the TCEQ describes their modeling choices and assumptions, and calculates an inhalation unit risk factor (URF), ultimately applying age-dependent adjustment factors (ADAFS) in DSD Section 3.4.2. Do you disagree with any of the modeling choices and assumptions or calculations made by TCEQ in the dose-response assessment? Please discuss any issues or concerns you have with the inhalation URF derivation.
10. Based on biomarker data, various sections of the DSD (e.g., Section 3.4.1.2.1, Section 3.4.1.4.2 number "4.", second to the last paragraph of Section 3.4.1.6.2) discuss air concentrations corresponding to endogenous and background EtO levels and also compare these levels to acceptable air concentrations derived from URFs (either the TCEQ's or EPA's). Such a discussion is also included in the Response to Public Comments document (e.g. Response to Dr. Kyle Steenland, Comment 3). Please comment on whether the information and context provided by the discussion of endogenous/background EtO levels is clear and is scientifically appropriate.
11. Please provide comments on the overall accuracy, objectivity, and transparency of the presentation of information in the revised DSD. Are the assumptions, data, and analyses described completely and clearly? Please identify any sections that need revision or improvement and describe in detail, to the extent possible, how they should be revised.
12. The TCEQ solicited public comments on a June 2019 proposed DSD and has prepared a response to those comments (See Response to Public Comments Received on the Ethylene Oxide Draft Development Support Document, January 2020). Has the TCEQ appropriately addressed the critical scientific questions and issues raised by the public commenters in the Response to Comments and/or revised DSD? Are the responses to public comments presented clearly and completely? Please explain.
13. Please discuss any additional relevant comments or issues. Are there any additional questions or concerns that you would like fellow peer reviewers to address?

Clarifying Questions on Draft Responses to Charge Questions

TCEQ and the experts reviewed the draft responses to the charge questions. TCEQ had a number of questions seeking more information and/or clarification from the experts regarding their draft responses. In addition, an expert requested input on a specific topic from other experts. The peer reviewers verbally responded to the clarifying questions during their April 8, 2020 teleconference with TCEQ and have revised their responses to the charge questions as they felt it was appropriate. Below is a list of the clarifying questions, organized by relevant charge question.

TCEQ Clarifying Questions

- Charge Question 3: Could you provide information on mechanisms of action that can lead to a flattening of the risk curve as a function of exposure, with applicable references?
- Charge Question 3: Please provide references or further methodology details regarding the outline of a suggested method for informing the dose-response curve with ancillary data and potential use of a Bayesian approach.
- Charge Question 4: Please expand upon the factors you considered in deciding whether there is a causal relationship between EtO and breast cancer. Did the positive studies have qualities that allowed them to be more heavily weighted than the negative studies? How did you interpret and take into consideration the width of the confidence intervals around the risk estimates?
- Charge Question 4: When assessing the weight of the epidemiological evidence, how would you choose between assessing the results from continuous versus categorical models? How would you choose between external versus internal referent analyses?
- Charge Question 5: Does using the 95% upper confidence of the slope to calculate risk estimates help to account for the difference between mortality and incidence for lymphoid cancers?
- Charge Question 6: Please elaborate on how to pick amongst models if two (or more) models fit the data equally. Do you choose or prioritize based on protectiveness, prediction of observed data, simplicity, alignment with biological processes, and/or other factors?
- Charge Question 8: Which data were considered “low dose data” in comments regarding the fidelity of model fit with low dose data?
- Charge Question 13: Could you provide further information about what aspects or sections of the document are referred to in reference to “overreliance on statistical significance in judging the epidemiologic evidence”?

Expert Question for Other Peer Reviewers

- In relation to Charge Question 8, an expert stated: “I am curious to see what epidemiology experts have to say about the purported bounds on predicted numbers of cases.”

After the teleconference, TCEQ submitted two additional clarifying questions, which were sent to the experts for consideration in revising their responses to the charge questions. The first TCEQ question requested clarification of a specific sentence in an expert’s draft response. The expert reviewer chose to delete the sentence from that expert’s final response to comments. The second question is listed below.

- Charge Question 9: Please clarify what you mean by comparing the analysis from section 3.4.1.5.2.2 (which discusses the breast cancer model and URF) with the estimates for the lymphoid cancers.

Expert Peer Reviewer Responses to Charge Questions

RSC randomly assigned each reviewer a number and used that number to identify the reviewer's responses to the charge questions. Responses have been collated by charge question below. In the responses, quoted passages from the US EPA and TCEQ documents have been italicized. Yellow highlights are experts' emphasis. Expert 2 also provided editorial annotations to a copy of the draft DSD document. A copy of this document has been provided to TCEQ.

Charge Question 1

The TCEQ conducted a systematic review of the literature relevant to the derivation of an inhalation unit risk factor for ethylene oxide (EtO) (see DSD Appendix 1). Are you aware of any additional literature or studies that should be considered and if so, how might they impact the assessment?

Expert 1

Yes, I have sent copies of 16 relevant peer-reviewed publications to Ms. Patterson who distributed them to the review committee (see list in Appendix D). Of these, findings from the following article not reviewed would add valuable information to the TCEQ review. See response to Q.4 for details.

Wong O and Trent LS (1993) Brit J Ind Med 50:308-16.

Expert 2

To this reviewer (not expert in systematic review), the review appears to be done appropriately. See specific comments below.

Appendix 1. Table 21. [Describing types of studies excluded from further consideration]

| | |
|------------------------------------|--|
| <i>Relevant endpoints examined</i> | <i>- Study focused solely on cytotoxicity - Study only measured sister chromatid exchanges (SECs), protein adducts, or chromosomal changes</i> |
|------------------------------------|--|

In my review I remarked on the inadequate evaluation of studies in support of or contradicting a mutagenic MOA for EtO. If TCEQ does an independent assessment of the MOA, then studies of the sort in column 2 will need to be retrieved and evaluated.

p.94 - "2. Sufficient human data exist for EtO such that animal data, although used to strengthen the carcinogenicity class, would not be used to derive a chronic carcinogenic toxicity factor. TCEQ (2015) states that in general, human data are preferred over animal data when developing toxicity factors."

At other points in the DSD, the human data are characterized as less than sufficient.

p. 100 - "Each reviewer scored the included studies independently, then came together to agree on a single score for each domain/study (individual scoring not shown)."

Whence came the reviewers? Were these TCEQ experts, or was the review done in some other way?

p. 104 - "5. Unlike USEPA (2016) that uses a lifetime exposure duration value of 85 years, the TCEQ-directed dose-response analyses use a standard default of 70 years consistent with TCEQ guidance (TCEQ 2015)."

Apart from policy, is there an advantage to using 70 years as lifetime? Is this closer to lifetime in the study population or in the population likely to be affected by the risk assessment?

[Note added after review of the TCEQ Response to Comment (RTC) document: In the RTC there is some discussion of the use of 70 years vs. 85 years in various contexts (life tables, duration of exposure). I suggest that this discussion in the RTC could be better reflected in the DSD in Appendix 1 and in the summary chapter.]

Expert 3

No, I am not aware of additional literature or studies to be considered, but I have not conducted an independent literature search for such articles. The search strategy used by TCEQ is very likely to have identified all relevant literature. Based on my general familiarity with the epidemiologic literature, it seems quite plausible that the volume of original studies is limited and that all the major ones likely to contribute to a quantitative risk assessment of this nature have been incorporated. However, since there are a significant number of epidemiologic studies that have been completed, a clear statement of why the only two that were considered in detail were the NIOSH and Union Carbide studies. I expect that is because those are the only ones that provide quantitative exposure estimates, but regardless, a statement on the choice to focus on those is warranted to make clear that all potentially contributory articles were given due consideration.

Expert 4

No. I'm not aware of additional literature or studies that should be considered (relevant to EtO); in fact, it appears that the database has not evolved substantively since the conduct of earlier assessments (e.g., US EPA, 2016), with more recent references being based principally on the reanalysis of preexisting information. While additional information directly relevant to the assessment of EtO – i.e., epidemiological, toxicological and/or mechanistic data have not been identified, references relevant to the methodological approach are cited in several of the responses to other questions (e.g., Questions 2,4,6).

Additional references that may be relevant to consideration of the shape of the dose-response curve at low dose (question 3), are potentially those referenced in Figure 12 of Vincent et al. (Dose-Response: An International Journal October-December 2019:1-17). These references are not cited in the article but appear to be based on those included in Tables 3-6, 3-7 and 3-8 of the U.S. EPA assessment. The form of this analysis is helpful in analyzing data on dose-response by key events in a hypothesized mutagenic mode of action. However, what's missing in the Vincent et al. reference is more detailed analysis of the incidence data across key events in a graphical presentation based on supporting tabular concordance tables. Focus on early key events might be informative in relation to the shape of the dose-response curves. While I've not conducted a literature review, some of the references identified by Reviewer 5 in relation to question 3 may additionally inform such analysis (in particular, that of Manjanatha et al. [2017]). Of particular interest, also, in any analysis of this type would be observation of earlier key events (i.e., consistent with a mutagenic mode of action) in human populations (reviewed recently in the IARC monograph though there appear to be few very recent studies (<https://monographs.iarc.fr/wp-content/uploads/2018/06/mono100F-28.pdf>)).

Expert 5

I have provided a list of potential literature citations in my response to Charge Question 3. These sources relate to mutagenesis studies, among other topics related to estimating low-dose dose-response patterns. I believe such studies were excluded from consideration at the time of the literature review. However, I suggest that these and other sources discussed in response to Charge Question 3 could be useful for informing an alternative approach to estimating the low-dose shape of the EtO dose-response curve.

Expert 6

I am not aware of any additional relevant literature, although I am not familiar with the ethylene oxide literature in general.

Charge Question 2

The TCEQ adopts the EPA conclusion that the weight of the evidence supports a direct-acting mutagenic mode of action (MOA) for EtO carcinogenicity (DSD Sections 3.3 and 3.3.1). Section 3.3.1 of the DSD presents summary information from the EPA (EPA Section 3.4.3) relevant to the MOA determination. Do you agree with the MOA determination? Please explain.

Expert 1

Outside my area of expertise

Expert 2

I can neither agree nor disagree with the chosen MOA of mutagenicity as the description of MOA evaluation is inadequate. US EPA did not provide a complete and convincing MOA evaluation, and the TCEQ provided only a cursory, not entirely accurate description of the EPA evaluation. I strongly suggest that TCEQ do an independent assessment of EtO MOA using the US EPA framework (which US EPA did not do in a comprehensive fashion) or some other contemporary evaluation process. The TCEQ document must be convincing on its own. It is not.

Given that TCEQ has chosen to accept some of the US EPA conclusions on EtO and to refute others, it is absolutely necessary that the rationales for acceptance or rejection be clear and explicit. The TCEQ document should be specific as to why they chose to accept US EPA's MOA, as they later in the document choose to depart from the EPA's model choice. The rationales for both choices need to be explicit. Rather than accepting a poorly described MOA in the EPA document, I strongly recommend that TCEQ conduct a contemporary evaluation of the EtO MOA. The choice may still be a MOA with mutagenicity by EtO as an early key event, but the support ought to be explicitly described, and the areas wherein data are weak or not supportive described as well.

In my opinion US EPA did not do a thorough evaluation, but rather gave the mutagenicity MOA a free pass: the agent seems to be mutagenic, so the MOA is mutagenicity. This does not meet contemporary standards. The burden of proof for demonstrating that an agent has a mutagenic MOA is as high as that for cytotoxicity, receptor binding or any other MOA. In their document, US EPA did not provide sufficient descriptions of empirical support for mutagenicity as an early key event in the carcinogenic process. Their descriptions were particularly deficient in the dose and time concordance aspect of the MOA support.

In the US EPA (and subsequently in the TCEQ) documents, there appears to be no judgement of study quality for the genotoxicity database. There seems to be a universal acceptance by US EPA of all positive genotoxicity findings with no consideration of the quality of the data or whether studies comport with contemporary guidelines for assay conduct and interpretation.

US EPA does not provide a temporality / dose response matrix, and does not adequately document either the temporality for early key events (for which there appear to be data), or the expected observation of early key events at lower dose than later key events (for which there may not be data to support the MOA). The point here is not that there is a dose response for some observation related to a key event, but rather that hypothesized early key events are observed at exposure levels lower than those associated with late key events, such as cancer. The US EPA document says "*Mutation frequency in the reporter genes Hprt and LacI was increased in a concentration-dependent manner primarily in lymphocytes from rats and mice exposed to concentrations associated with significant tumor induction in cancer bioassays (i.e., ≥ 50 ppm) for up to 48 weeks,*" (emphasis added) p. 3-55. This is not strong dose response support for mutation as an early key event.

Later in their document US EPA uses highly questionable language that I recommend TCEQ not copy. Some examples.

"EtO-induced genotoxicity is observed after shorter exposure durations and at lower exposure concentrations than those associated with tumor induction in both rodents and occupationally exposed humans (see Section 3.3.3.4)." p 3- . This statement is not substantiated by the presentation of a key event dose /time matrix. It appears from the cited section that there may be some temporal concordance, but mutation generally is observed at only the higher tested doses.

". . . and there is incontrovertible evidence that EtO is mutagenic and genotoxic (see Sections 3.3.3.2, 3.3.3.3, and 3.3.3.4)." I find this to be an unacceptable overstatement. Much of the data is from older studies, which are unlikely to have been conducted under contemporary guidelines (particularly for interpretation of responses). Moreover, non-positive studies are not discussed. And it appears that all positive results are considered to be equally indicative of strong support for mutation by EtO as an early key event.

Data showing differences in critical gene mutational spectra between tumors in EtO treated and control animals do provide some support for the consistency criterion for a mutagenic MOA. But these data do not provide strong support that mutation is an early key event; mutations continue to accumulate in neoplastic cells throughout the carcinogenic process.

Description of clastogenicity and other indications of large DNA changes (such as SCE) belong in the MOA description. These observations, while they may be supportive of mutation as an early key event, do not provide strong support for low dose linearity, and certainly do not provide a rationale for supralinearity.

The DSD should specify that in this document the definition of EtO as direct acting means that no metabolism is required for Sn2 reaction. I suggest that DSD copy the language below from US EPA (2016).

EPA p 3-29 - "EtO is a direct-acting SN2 (substitution-nucleophilic-bimolecular)-type monofunctional alkylating agent that forms adducts with cellular macromolecules such as proteins (e.g., hemoglobin, see Section 3.3.2) and DNA"

Please see a number of specific comments below.

3.3.1 MOA, p.19 par 4 - "The hypothesis is that EtO carcinogenicity has a mutagenic MOA. This hypothesized MOA is presumed to apply to all the tumor types. The key events in the hypothesized mutagenic MOA are: (1) DNA adduct formation by EtO, which is a direct-acting alkylating agent; (2) the resulting heritable genetic damage, including DNA mutations, particularly in oncogenes and tumor suppressor genes, as well as chromosomal alterations; and (3) the clonal expansion of mutated cells during later stages of cancer development; eventually resulting in (4) tumor formation. Mutagenicity is a well-established cause of carcinogenicity."

This is rather inadequate. Mutations are heritable changes. Some alkylations are heritable but do not result in changes in gene expression. Note that mutation must be in cancer critical genes for the agent to cause cancer by a mutagenic MOA.

Who says that mutagenicity is a cause of cancer? There should be a citation. This overly general statement implies that mutagens are carcinogens, which is not the case. All cancers have mutations associated with them, but not all carcinogens cause cancer by virtue of their ability to cause mutations in cancer critical genes as an early event in the process.

p. 19 last par - "Numerous studies have demonstrated that EtO forms protein and DNA adducts, in mice and rats, and there is incontrovertible evidence that EtO is mutagenic and genotoxic."

Genotoxic is not mutagenic. Who says that the evidence is incontrovertible? There should be at least some examples provided of positive studies that meet contemporary standards.

"The evidence for causal associations between the key events and tumor formation has strength and consistency. Increases in the frequency of gene mutations in reporter genes have been observed in the lung, T-lymphocytes, bone marrow, and testes of transgenic mice and in T-lymphocytes of rats exposed to EtO via inhalation at concentrations similar to those inducing tumors in the rodent carcinogenesis bioassays."

Note that the doses for mutation observation should be lower than tumor inducing doses. Dose time matrix should be provided.

p. 20 par 2 - "In rats, although SCEs are consistently observed in the available studies, the results for micronuclei formation and chromosomal aberrations following subchronic (up to 4-week) inhalation exposures to EtO at the same exposure levels as those used in the rodent bioassays have been nonpositive; however, IARC (2008) has noted analytical limitations with some of these analyses."

SCE are reflective of interaction with DNA rather than mutation. There is much more evidence for correlation of observations of both mutations and positive cancer bioassays with micronuclei and chromosome aberrations. The point regarding study quality is raised here, but it should be made throughout.

"In addition, Donner et al. (2010) demonstrated a clear duration effect in mice, with chromosomal aberrations being induced at those same EtO exposure levels only following longer exposure durations (≥ 12 weeks)."

Not sure what argument is being made here. This doesn't look like a dose / time matrix observation.

p. 20 par 3 - "A temporal relationship is clearly evident, with DNA adducts, point mutations, and chromosomal effects observed in acute and subchronic assays."

So why is there no dose time matrix? This is not convincing.

p. 20 par 4 - “Dose-response relationships have been observed between EtO exposure in vivo and DNA adducts, SCEs, and Hprt and Trp53 mutations.”

The point for demonstrating a MOA is not that any particular observation has a dose response with EtO, but rather that hypothesized early events are observed at lower doses and at earlier times than later key events.

p. 20 par 5 - “In addition to the clear evidence supporting a mutagenic MOA in test animals, there are no other compelling hypothesized MOAs for EtO carcinogenicity. For example, there is no evidence of cytotoxicity or other cellular dysfunction indicative of regenerative proliferation, and little-to-no evidence supporting some other toxicity-related MOA, such as oxidative stress.”

Absence of proof is not proof of absence. What is being described here? There is no effort in the DSD or US EPA to describe other MOA. Observed mutagenicity for the agent does not negate the likelihood of other MOA as well, or instead.

p. 21 par 1 - “DNA adducts in EtO-exposed humans have not been well studied, and the evidence of increased DNA adducts is limited. EtO has yielded positive results in in vitro mutagenicity studies of human cells.”

At relevant levels of exposure? At various points in the US EPA document it was noted that genotoxic effects were observed in the range of exposures that caused tumors. This observation does not support a conclusion that EtO caused mutations as an early key event.

“Although the studies of point mutations in EtO-exposed humans are few and insensitive and the evidence for mutations is limited, there is clear evidence from a number of human studies that EtO causes chromosomal aberrations, SCEs, and micronucleus formation in peripheral blood lymphocytes, with some evidence of positive relationships with exposure concentration and duration.”

Again, are the levels relevant to MOA for cancer?

Expert 3

This is outside the scope of my expertise.

Expert 4

I agree that the available data support a mutagenic mode of action. However, a more systematic analysis of the extent of the supporting evidence for the hypothesized mutagenic mode of action based on consideration of the essentiality and empirical support for the hypothesized Key Events (KEs) and Key Event Relationships (KERs) would have provided stronger rationale. For example, developing a concordance table – i.e., tabulating incidence data for each of the key events at increasing doses across various durations of exposure transparently characterizes the extent to which evidence is consistent with expected patterns of empirical support. This format also identifies dose-response relationships for key events (including those in humans) that may be most helpful in “ground truthing” the estimated cancer risks from the epidemiological studies. Vincent et al. (2019) have made some progress in aligning such information in Figure 12 of their publication; however, presentation here falls short of that which is most informative (i.e., including the quantitative incidence data).

Recent developments in more systematic definition and application of the modified Bradford Hill considerations for mode of action and adverse outcome (AOP) pathway analysis are documented in the following references; see, for example,

Meek et al., 2014a <https://onlinelibrary.wiley.com/doi/pdf/10.1002/jat.2949>;

Meek et al., 2014b <https://onlinelibrary.wiley.com/doi/full/10.1002/jat.2984>;

The OECD AOP Handbook: <https://www.oecd-ilibrary.org/docserver/5jlv1m9d1g32-en.pdf?expires=1584027129&id=id&accname=guest&checksum=A8F93325CADE89A776FE6799447D5E4A>.

In the mode of action analysis cited in the DSD (i.e., that of US EPA), dose-response relationships for each of the KEs and temporality are considered separately, rather than based on assessment of their joint concordance across the KERs.

Expert 5

Based on the evidence presented, I am satisfied with the MOA determination. My understanding, *a priori*, was that EtO was a direct-acting mutagen. I saw no evidence to suggest that that was inappropriate.

Expert 6

A direct-acting mutagenic mode of action (MOA) for EtO carcinogenicity seem reasonable in light of evidence that ETO can induce a wide range of genetic damage. However, I am not an expert in this area.

However, there is a comment by TCEQ in response to University of California, San Francisco Comment 2 that *“The DSD quite plainly states that as USEPA acknowledges, MOA data do not support their model.”* I don’t know what this statement refers to but all I can think of is that it relates to TCEQ’s classification of the EPA model as a “supra-linear” whereas a mutagenic MOA suggests the dose response is (low-dose) linear. If that is what this comment refers to, then it is clearly wrong because the EPA dose response is low-dose linear, in fact, it is exactly linear in the low-dose range. (For a more complete discussion of this point see my response to question 7.)

Charge Question 3

The TCEQ adopts EPA’s MOA analysis (DSD Section 3.3.1) and considers MOA as information relevant to the likely or expected shape of the dose-response (DSD Sections 3.4.1 and 3.4.1.1) as specified by the TCEQ guidelines for developing toxicity factors (TCEQ, 2015). What is your opinion on whether and how the MOA should inform the likely or expected shape of the dose-response curve, overall and in the low-dose range (e.g., at environmentally-relevant concentrations); and whether and how the MOA should inform the choice of dose-response model for estimating human carcinogenicity risk? Please comment on TCEQ’s reasoning on the implications of the MOA for the shape of the dose-response and its relative importance amongst their other model choice considerations (summarized in DSD Section 3.4.1.4.2). Are the TCEQ conclusions concerning implications of the MOA scientifically defensible?

Expert 1

Outside my area of expertise.

Expert 2

This reviewer considers US EPA demonstration of a mutagenic MOA for EtO to be inadequate as presented in US EPA (2016) and described in the DSD. Please see comments above.

If no MOA has been demonstrated, then US EPA (2005) accepts as a default the use of linear low dose extrapolation. Application of defaults (linear or less than linear extrapolation procedures) are used when there are insufficient data or models to apply a biologically based dose response model (US EPA 2005). I consider these to be reasonable policy choices.

When there has been a demonstrated MOA for a chemical this information can be appropriately applied to the choice of low dose extrapolation models. While linearity at low dose has often been used for modeling cancer risk when a mutagenic MOA has been specified, it is acknowledged by scientists in the field that mutation is not a single step process. The lack of linearity at low dose has been demonstrated for some observations of mutagenicity for a number of potent carcinogens (e.g. some alkylating agents). Large deletions, transversions, clastogenicity, and some other mutagenic effects are generally considered to be non-linear at low dose. DNA adduct formation by some agents appears to have a linear low dose slope; this is not the case for other agents. One of the advantages of MOA articulation is that it can permit the choice of an observed early key event as the point of departure. I suggest that the DSD discuss not only the choice of the critical study population, but also the effect (cancer mortality vs. some early key event) to be used in dose response modeling. Presumably the cancer mortality data are the only ones that would support modeling, but it would be useful to clarify this in the document.

Description of clastogenicity and other indications of large DNA mutations belong in the MOA description. These observations, while they may be supportive of mutation as an early key event, do not provide strong support for low dose linearity, and certainly do not provide a rationale for supralinearity.

In short, MOA (properly described and supported) can and should be used in defining choices for low dose extrapolation and / or application of models.

Please see some specific comments below.

3.4.1.1 Consideration of MOA - p. 26 par 1 - "the expected dose-response could be characterized as appearing sublinear in the low-dose range and/or sublinear overall across doses (see Figure 1)."

Or it could be two linear dose responses, with the lower dose slope more shallow than that of the higher doses (when repair or metabolism has been overwhelmed).

3.4.1.2.2 Key Data from the NIOSH Cohort and Endogenous Data - p. 34 last par - "certain critical cancer endpoints in the NIOSH cohort (i.e., all hematopoietic, lymphoid, non-Hodgkin's lymphoma) were only statistically increased in males, while breast cancer incidence was only statistically increased in females, and only in the highest EtO exposure quantiles for each of these cancer endpoints."

Given the relative rarity of breast cancer in males, is this observation unexpected?

Model choice - P64 par 2 - "MOA (i.e., the Cox proportional hazards model is indistinguishable from linear across doses of interest and appropriate for dose-response assessment of a direct-acting mutagenic carcinogen, particularly in the acknowledged absence of mechanistic data supporting an overall supra-linear dose-response; see Section 3.4.1.1);"

Need to be more explicit. Say either that (after the recommended re-evaluation) a mutagenic MOA does not contradict use of model. Or (if TCEQ finds insufficient support for a mutagenic MOA), linear is an acceptable default.

Expert 3

This is outside my range of expertise, but as one of multiple sources of input for making an informed judgment (which is all that a quantitative risk assessment can be, ultimately), it seems appropriate to use the information conveyed by insight into the mode of action. In fact, the epidemiologic data could not possibly provide the desired quantification of risk of cancer in the low-dose range so reliance on inferences based on mechanisms is required.

Expert 4

Firstly, as per the response to Question 6, in my view, mode of action is the primordial consideration in informing choices about extrapolation to the low dose region. However, it is often not sufficient in its own right and as a result, science policy choices are invoked. As per the responses to questions 1 and 2, more robust consideration of the concordance of dose-response relationships for key events in the hypothesized mode of action might additionally inform the nature of the dose-response curves at lower concentrations than those in the epidemiological studies of cancer incidence and mortality. It's also helpful in identifying biological "tipping points" (beyond which effects are likely irreversible). Conduct of such analyses seems advised. It's possible, though that the doses at which early key events have been investigated following exposure to EtO may be too high to additionally inform consideration of the shape of the dose-response curve at the potentially much lower environmental levels.

In the absence of relevant information to inform extrapolation to the low dose region, a decision concerning selection of the relevant model is a function largely of Agency policy, rather than scientific judgment. It seems important, then, to be as explicit as possible on the extent of reliance on Agency policy and previous Agency precedent and guidance in this context (see response to Question 7).

In relation to considering the shape of the dose-response curve at environmentally relevant concentrations, the arguments in Section 3.4.1.4.2 on endogenous formation (point 4 on page 64) are not particularly convincing (See also response to Question 7).

Some additional "reality checks" or "ground truthing" have been suggested, here, as a basis to strengthen the rationale for the TCEQ conclusions regarding appropriate modelling to the low dose range (See response to Question 8).

Expert 5

I do believe that MOA information can and should be used to inform the shape of the dose-response curve, especially at low doses. The implications of that information about shape for the choice of model, however, is perhaps not as straight-forward as some might think. And, in particular, I believe that the verbiage used by TCEQ is actually unhelpful in that respect.

I am specifically referring to the labeling of the USEPA spline modeling as "supra-linear." The USEPA model is piece-wise linear. The term "supra-linear" should be reserved for those curves whose slopes continually increase (to infinity) as the independent variable (e.g., dose or exposure) approaches a particular value, in most cases zero. That was certainly the context of the quote from Crump and Allen (1985) that is on pp 24-25. In other words, I believe that TCEQ has misconstrued the meaning of supra-linear and has cited Crump and Allen (1985) erroneously in its attempt to "disqualify" particular curve shapes, ones that are not in fact supra-linear.

What TCEQ appears to be objecting to is a curve that is steeper at low doses than at higher doses. That objection seems disingenuous. Later in the DSD (Section 3.4.1.3, p. 57) the TCEQ cites the Michaelis-Menten model as an example of a model that has a biological or mechanistic basis. But the Michaelis-Menten model has exactly the curve shape that TCEQ apparently finds objectionable; it is steeper (though not infinitely steep) as dose approaches zero, becoming progressively less steep as dose increases. The piece-wise linear model that USEPA applied to EtO is steeper (but not infinitely steep) at lower doses and less steep at higher doses.

I do not know if a Michaelis-Menten-type curve shape, specifically, would be mechanistically appropriate for modeling EtO-induced carcinogenicity. However, one might hypothesize that a direct-acting carcinogen such as EtO might display dose-response behavior like a receptor mediated response, the mechanism underlying the Michaelis-Menten model. In the case of EtO, the “receptor sites” would be the particular sites of mutation in the particular types of cells that would lead to a particular type of tumor (lymphoid tumors for example). Those “sites” would tend to get and remain “occupied” (i.e., mutated) as EtO concentration increased. The resulting pattern would be a flattening out of the risk as a function of exposure. Such reasoning, coupled with the assumed direct-acting mutagenic MOA, would tend to support a model that was essentially linear at low doses and that flattened out at higher doses. The USEPA spline models do that.

Nevertheless, as stated by TCEQ on p. 57, many models used for dose-response analysis and for risk assessment purposes do not have known biological or mechanistic motivations. Contrary to what is stated on p. 57¹ of the DSD, however, that would appear to make fit to the data more important, not less important, to the determination of model adequacy, as long as the models under consideration do not clearly violate conceptions of MOA. As presented above, it is my belief that the USEPA spline models do not clearly violate the MOA assumed for EtO. My comments related to model fit are presented in response to question 6.

In conclusion, TCEQ should expunge the incorrect characterization of the USEPA method as supra-linear, not erroneously cite Crump and Allen (1985) as support against the USEPA spline models, and recognize that the MOA data that have been used by TCEQ to criticize the USEPA model are much less powerful for model selection than they appear to think.

Having said all that, however, the real question is, should a (presumably) low-dose linear part of the dose-response curve be estimated from observations that are associated with much greater exposures. That is what USEPA has done with their spline models. I do not think the DSD has presented a focused and consistent rebuttal to that.

In fact, consider the following rather back-of-the-envelope calculation that stands in contrast to that provided as a reality check in the DSD pp. 30- 31 (Section 3.4.1.2.1.1). As done in the DSD let’s focus on lymphoid cancers. In fact, let’s consider lymphoid cancers other than AML, which is, according to Kirman and Hays (2017), the only lymphohematopoietic cancer linked causally to smoking. That means that we can use the overall background rate for all other lymphohematopoietic cancers as the rate that applies to nonsmokers. My approximate calculation is that the appropriate background rate to use is 0.025 (3% for all lymphoid cancers as cited in the DSD, minus 0.5% for AML suggested by a quick on-line search²).

¹ The DSD states: “Thus, in this respect model fit alone is a lesser consideration compared to data (e.g., MOA data) that may (or may not) adequately support use of a particular model.” I find this sentence a bit confusing, following as it does the statement that many models used for risk assessment lack a biological or mechanistic basis.

² See <https://www.cancer.org/content/dam/CRC/PDF/Public/8674.00.pdf>, p. 4.

Ignore, for this exercise, exogenous EtO exposure; consider only endogenous EtO and the corresponding air concentrations that have been derived by Kirman and Hays (2017) (Table 4, for non-smokers). Let e be the corresponding air concentration (ppb) for any given endogenous concentration. Kirman and Hays (2017) derive a distribution for e for non-smokers (Table 4), let's call it $p(e)$. Then,

$$\text{background lifetime cancer probability} = \int_0^{\infty} p(c|e) \cdot p(e) de$$

where $p(c|e)$ is the probability of cancer given an endogenous concentration corresponding to e .

From TCEQ's MOA argument, the effect of endogenous EtO should be the same as the effect of exogenous EtO and should have a linear relationship:

$$p(c|e) = a + b \cdot e$$

in the range of values of e such that $p(c|e)$ is relatively small. For the sake of illustration, let us suppose that a high percentage, 90%, of the lymphoid cancers under consideration have nothing to do with EtO. If that is true, then $a = 0.9 * 0.025 = 0.0225$.

Then,

$$\begin{aligned} \text{background lifetime cancer probability} &= \int_0^{\infty} (a + b \cdot e) \cdot p(e) de \\ &= a + b * E(e) \end{aligned}$$

where $E(e)$ is the expected value of e , in this case for non-smokers. Let us estimate $E(e)$ by the mean value from Table 4 of Kirman and Hays (2017), 1.9 ppb. Plugging the numbers into the equation we get

$$b = (0.025 - 0.0225)/1.9 = 0.0013 \text{ per ppb}$$

Note that this estimate is akin to a URF, and that it is only about a factor of 5 less than the URF promulgated by USEPA (0.0071 per ppb), which has been made conservative by considering an upper bound on the slope parameter and by applying an ADAF, neither of which has been done in this quickish calculation.

Also note that, as far as I know, the percentage of non-AML lymphoid cancers that have absolutely nothing to do with EtO exposure is not known, and may be unknowable. And TCEQ makes no claims that that percentage should be greater than 90%; the DSD does not even consider that question because, in my opinion, its reality check is flawed. Moreover, if all lymphoid cancers (minus AML) are in some way (roughly linearly) related to (affected by) EtO exposure, then the estimate of b becomes 0.013 per ppb, a value greater than the URF promulgated by USEPA.

If the above calculations are close to being correct, then one must conclude that the reality check considering endogenous production of EtO does not argue at all strongly against the health-protective USEPA slope estimate (URF).

My conclusion is that TCEQ has *not* argued coherently that the EPA assumption (that the truly low-dose linear slope is well-estimated from higher-exposure data) is incorrect. The questions then become: what evidence or data would contribute to TCEQ's case, would the revised truly low-dose slope be different (presumably less) than estimated by USEPA, by how much, and in what ranges of exposure?

In that regard, I think that TCEQ (and USEPA) have missed an opportunity to utilize information from "ancillary" studies to assist with that determination. For example, data on sister chromatid exchanges,

protein adducts, or chromosomal changes (among others, perhaps) provide direct evidence of events potentially related to mutagenic activity that may be observed at lower levels of exposure. Note, however, that TCEQ has explicitly excluded consideration of such data, based on their literature search and exclusion criteria (see DSD Table 21).

My conception of a possible path forward is outlined as follows:

- Propose low-dose linear models for EtO carcinogenicity. The models need *not* postulate a slope that is the same (constant) from 0 dose all the way to 1600 ppm-days (as the USEPA-chosen spline model does). In fact, the models should probably have an inflection (or transition) point at which the truly low-dose slope changes to greater slope(s) as the biological conditions dictate. The models may have a subsequent flattening of the dose-response at even higher doses.
- Consider ancillary data to inform the estimate of the truly low-dose slope. I conceive of this being a Bayesian analysis, where the ancillary information can be used to define priors for the parameters of the cancer dose-response model, including if need be the location of the aforementioned inflection (or transition) point.
- Use the epidemiological data to update the model parameters and compute posterior distributions for doses associated with risk levels of interest (e.g., 1e-5). The credible interval estimates from that process could then be used to define a health-protective exposure level (e.g., using the 95% credible interval lower bound).

With respect to the Bayesian application, I have no specific references to such approaches being applied to the derivation of toxicity factors. However, there have been some recent Bayesian applications to epidemiological data, and dose-response analyses in particular. Moreover, several authors have pursued the idea of concordance of dose-response patterns between precursor events and apical endpoints. A Bayesian implementation of the ideas put forward in those papers might include the use of that information to define priors related to the truly low-dose response pattern. In fact, there are publications specifically about the mutagenic potential of EtO that may prove to be useful. A bibliography of references related to all of these topics is included here:

Allen, B., Shao, K., Hobbie, K., Mendez, W., Lee, J., Cote, I., Druwe, I., Gift, J., Davis, J. (2019). Bayesian Hierarchical Meta-Regression of Epidemiological Studies, Part 2: Dose-Response Modeling and Target Population Predictions. Submitted.

Vincent, M. J., Allen, B., Palacios, O. M., Haber, L. T., & Maki, K. C. (2019). Meta-regression analysis of the effects of dietary cholesterol intake on LDL and HDL cholesterol. *The American journal of clinical nutrition*, 109(1), 7-16.

Manjanatha, M., Shelton, S., Chen, Y., Parsons, B., Myers, M., Gollapudi, B., Moore, N., Haber, L., Allen, B., and Moore, M. (2017). Dose and temporal evaluation of ethylene oxide-induced mutagenicity in the lungs of male big blue mice following inhalation exposure to carcinogenic concentrations. *Environmental and Molecular Mutagenesis*, 58(3), pp.122-134.

Allen, B., Vincent, M., Liska, D., and Haber, L. (2016). Meta-regression analysis of the effect of *trans* fatty acids on low-density lipoprotein cholesterol. *Food and Chemical Toxicology* 98(Pt B):295-307. DOI: 10.1016/j.fct.2016.10.014.

Manjanatha, M.G., S. D. Shelton, L.T. Haber, B. Gollapudi, J.A. MacGregor, and M.M. Moore. 2015. Evaluation of *cII* mutations in lung of male Big Blue mice exposed to vanadium pentoxide by inhalation for up to 8 weeks. *Mutat Res Genet Toxicol Environ Mutagen*. 789-790:46-52.

Banda, M., K.L. McKim, L.T. Haber, J.A. MacGregor, B.B. Gollapudi, and B.L. Parson. 2015. Quantification of *Kras* mutant fraction in the lung DNA of mice exposed to aerosolized particulate vanadium pentoxide by inhalation. *Mutat Res Genet Toxicol Environ Mutagen*. 789-790:53-60.

Parsons BL, Manjanatha MG, Myers MB, McKim KL, Shelton SD, Wang Y, Gollapudi BB, Moore NP, Haber LT, Moore MM. 2013. Temporal changes in K-ras mutant fraction in lung tissue of big blue B6C3F₁ mice exposed to ethylene oxide. *Toxicol Sci*. 2013 Nov; 136(1):26-38. doi: 10.1093/toxsci/kft190

Moore, M., Heflich, R., Haber, L., Allen, B., Shipp, A., and Kodell, R. (2008). Analysis of in vivo mutation data can inform cancer risk assessment. *Regulatory Toxicology and Pharmacology* 51:151.

Allen, B., Zeiger, E., Lawrence, G., Friedman, M., and Shipp, A. (2005). Dose-response modeling of in vivo genotoxicity data for use in risk assessment: some approaches illustrated by an analysis of acrylamide. *Reg. Tox. Pharm.* 41:6-27.

Expert 6

I consider that it is reasonable to assume that a chemical such as ETO, which can induce a wide range of genetic damage, has a dose-response for human carcinogenicity that is linear at low dose. However, this fact alone does not provide any basis for deciding between the Cox proportional hazards model developed by TCEQ and the spline model developed by USEPA, because they both have this property. This issue is made more confusing by the TCEQ characterizing the USEPA model as “supra-linear” (see discussion on question 7 below).

Charge Question 4

The TCEQ conducted an evaluation of EtO’s carcinogenic classification (DSD Section 3.3.2), and also evaluated breast cancer risk in humans as a potential cancer endpoint (DSD Appendix 6; Response to Dr. Kyle Steenland, Comment 1 in Response to Public Comments Document). What is your characterization of the overall weight of the evidence for or against EtO increasing the risk of breast cancer in humans at occupational concentrations (past or present) and at environmentally-relevant concentrations?

Expert 1

Insofar as the workplace typically represents the “high dose” setting, epidemiologic investigation of cancer risks in worker populations generate important data for etiologic inference. Insofar as hazardous occupational exposures typically have been reduced over time due to worker safety policies, occupational epidemiology studies can address changes in risk related to exposure reductions. Whether this generalization applies to EtO exposure needs to be characterized. There has been limited occupational epidemiology research on EtO exposure and breast cancer risks. I identified only 5 relevant peer-reviews publications. Steenland et al. (2003) reported a dose-response trend for EtO and breast cancer incidence among a cohort of 7576 women workers in a US sterilization facility. The relative risk in the highest cumulative exposure quintile, lagged 15 years, and adjusted for reproductive risk factors was

1.87. The finding for breast cancer mortality in the highest cumulative exposure, lagged 20 years, was 3.13 (Steenland et al., 2004). Coggon et al.'s (2004) cohort mortality study of 1012 women EtO exposed workers at 3 UK plants indicated no association with breast cancer, as the Standardized Mortality Ratio (SMR) was 0.84. Mikoczky et al. (2011) conducted a cohort study among 1309 Swedish woman workers employed for at least 1 year in sterilization facilities. They reported a standardized incidence ratio for breast cancer of 3.55 among workers at the highest cumulative exposure level (>0.22 ppm-yrs). There was no evidence for an etiologic relation in a cohort mortality study of 10,019 women employed in 14 sterilization US facilities; the RR (95% CI) was 0.80 (Wong and Trent, 1993). My conclusion regarding an etiologic relation between EtO exposure and breast cancer risk is based on the quality of epidemiologic research in terms appropriate study designs, rigorous approaches applied for exposure assessment and statistical analysis methods for dose-response estimation, and control for potential confounders. The magnitude of relative estimates, and their statistical precision indicated by the width of confidence intervals, and shape and statistical precision of the exposure-response curve are important but not essential requirements. Statistically significant relative risks >2.0 in the highest exposure category and statistically significant exposure-response trends provide strong evidence for causation. I regard findings from internal analyses, such as defining a reference group with never exposed as more meaningful than findings based in external reference group, such as Standardized Mortality Ratio's (SMR's) because the internal reference analysis controls for potential confounding factors, such as demographic and life style factors (e.g., smoking, dietary factors) more effectively than analysis based on external reference group analyses. I never rely on the epidemiologic findings satisfying all of the Bradford Hill criteria for reaching etiologic conclusions, other than the essential temporal relation between exposure and disease occurrence. This is consistent with Bradford Hill's recommendation. On balance, I would conclude that the epidemiologic evidence for an association between occupational EtO exposure and breast cancer risk is supportive for an etiologic role of EtO. The findings have not been fully consistent, but the data obtained from well-conducted studies of established cohorts support a causal relation.

Expert 2

There are several weight of evidence (WOE) judgments made by TCEQ, which are needed for a complete effects assessment. These include WOE to determine a cancer classification or descriptor, such as "*carcinogenic for humans*". A second is the WOE in support of demonstrating that a hypothesized MOA applies to a specific agent; I have discussed this above. A third is the WOE supporting a choice of a critical endpoint and study suitable for quantitative dose response assessment. It appears from the RTC document that Dr. Steenland is referring to both WOE efforts number 1 and 3.

I am not certain how or in which WOE judgments TCEQ considered the breast cancer data after their evaluation. It is clear that TCEQ chose not to derive their URF from the breast cancer data. Their analyses as described in the main text and in Appendix 6 support this as a reasonable choice.

I am uncertain as to what is the basis for TCEQ adoption of the EtO descriptor "*carcinogenic for humans*." It is not clear as to whether TCEQ thought that the breast cancer data added support for this classification or were irrelevant. I recommend a more explicit description of what TCEQ included under the WOE for hazard identification or cancer classification. It is certainly acceptable to find that a data set adds to the WOE for hazard identification but is not suitable for quantitation. It seems that EPA considered the breast cancer data supportive of their WOE.

I suggest again that TCEQ do an independent evaluation of the conditions under which EtO is likely to pose a carcinogenic hazard to humans.

Specific comments follow.

Appendix 6 - p. 152 par 1 - "Breast cancer requires a more detailed weight of evidence evaluation"

I don't disagree, but this is another "who says?" statement. Is this in response to Public Comments?

p. 152 and on –

The whole section suffers from unclear and judgmental writing. Some edits are indicated on a markup of the text. I think the appropriate arguments are there, but they are buried in table footnotes, or are unnecessarily hyped. The impression I received from the writing was not that of a sound scientific basis for excluding the breast cancer data from quantitative dose response assessment.

A6.2 Healthy Worker Effect and Under-Ascertainment Considerations

p. 154 last par - ", particularly where the carcinogen operates via a mutagenic MOA (e.g., EtO)."

Finish the argument. I am assuming that the rationale is something along the lines of carcinogens with a demonstrated mutagenic MOA are likely to be multisite? Systemic?

p. 156 par - "Instead, Steenland et al (2003). indicate [emphasis added], "Because of the issue of under-ascertainment, we have emphasized internal exposure-response analyses in our study rather than the use of external referent population."

Is TCEQ adopting the conclusion from Steenland et al (2003) for Mikoczy et al (2011)?

A6.3 Relevance of Laboratory Animal Data

p.156 last par - "Reported results show that breast cancer is more frequently/commonly induced in laboratory animal species by these agents than in humans."

Implies that a particular type of chemical of chemical class is responsible for lack of site concordance. But "these agents" in this case are just those chemicals that made up the study group.

Here and in the main text I sense a confusion of WOE for carcinogenicity with WOE in support of critical effect choice.

"More telling is that while there is 47% overlap between agents that cause lymphoid and haematopoietic cancers in humans and animals, there is only 20% overlap between agents that have been shown to cause breast cancer in humans and animals (Table 21.7 of IARC 2019)."

Did the authors of IARC (2019) report that this was a significant difference or was biologically important / relevant? It seems rather in the next sentence that IARC dismisses tumor site concordance between animal and humans for all tumor types.

p. 157 par 2 - "animal data for EtO-induced cancers cannot be relied upon to identify cancer sites or otherwise predict EtO carcinogenic response in humans."

As noted before, the authors conflate WOE for classification as to carcinogenicity with WOE in support of a critical effect choice. I propose (and I think US EPA agrees) that animal data for whatever tumor site can provide data supporting a cancer classification that is relevant for human risk assessment. For most presumed carcinogens there are no positive human data, and hopefully this situation will obtain to a greater degree in the future. The TCEQ authors are (I think) making a different argument: that the

existing animal data are not useful in selecting a critical effect for quantitative dose response assessment.

Expert 3

As noted in the report, the evidence on EtO and breast cancer is limited in volume and aside from the NIOSH cohort, quite limited in terms of the quality of the epidemiologic evidence. To the extent that the Marsh et al. meta-analysis did not effectively distinguish studies based on quality, and it appears that they did not, the failure to observe associations in a number of lower-quality studies may well drive the aggregate risk estimate down. The meta-analysis is thus not very helpful since it does not address study quality. Therefore, the question about inclusion of breast cancer depends on the interpretation of the results from the NIOSH cohort.

Review of the incidence and mortality data on breast cancer in the NIOSH cohort reveals a notably mixed set of findings. There are some clear suggestions of an association being present but only with an extensive lag period and largely limited to the uppermost quartile. This could indeed reflect a causal effect that requires an extensive period before appearing (long latency) and is limited to the highest exposure group. But it could also reflect random error given the lack of overall excess and limited support for a dose-response gradient.

An additional complexity in studying occupational exposures and breast cancer is the role of parity which is strongly related to risk of breast cancer (higher parity predicts lower risk) and strongly related to remaining in the work force to accrue greater exposure (more live births predict cessation of employment). Without careful control in the analysis, this would result in a spurious positive association – the women with no or few children have elevated risk of breast cancer and work for longer periods of time, thus accruing greater cumulative exposure. In the breast cancer incidence study, parity was controlled in the analysis but without more detail on how this was done particularly in relation to the exposure lag periods, it is not clear that it was effectively handled. In other words, parity is a time-varying attribute as childbirth occurs but exposure is lagged so it is not clear if the parity applied to a given person-year at risk was that of the time at which cancer was occurring or the time during which the lagged exposure was occurring. The stated attempt to control parity mitigates the concern but does not fully remove concern. In fact, the finding that duration of exposure was more strongly associated with breast cancer incidence than cumulative exposure (duration times concentration) would be consistent with this hypothesized bias, i.e., working longer predicts higher risk.

Given inherent limitations noted in the NIOSH studies of both breast cancer incidence and mortality, suggestive but not compelling results limited to specific analytic approaches, and the absence of confirmatory data, an informed, unbiased evaluator could well come to the judgment that TCEQ did, i.e., not considering breast cancer in the overall EtO assessment. In fact, I agree with TCEQ's judgment even if my reasons are somewhat different. Although I am not entirely confident of the appropriateness of their judgment, in balance, I think that TCEQ made the right choice to not consider breast cancer in their quantitative assessment. Nonetheless it should be acknowledged that this is a judgment call and while their decision is fully defensible, arguments can be made for having made a different decision.

A question was raised regarding the interpretation of categorical and continuous results. Categorical results have the advantage of being more readily interpretable as ratios of the risk in higher versus lower exposure groups, e.g., a doubling of risk or increase by a factor of 1.3. Continuous measures require interpreting less familiar measures, e.g., risk increases by X% per unit of exposure.

Another question concerned interpretation of results based on comparisons of workers with the general population versus comparisons among workers with differing exposure levels. The healthy worker effect generally reduces the risk among working people compared to general populations, more for cardiovascular disease than for cancer, so the comparison of the health of workers to the general population is not necessarily indicative of the causal effect of workplace exposures. Comparisons of disease risk among workers with differing exposure histories is generally more informative regarding the impact of the exposures of interest since it is not distorted by the healthy worker effect.

Expert 4

TCEQ indicates that “Specifically, the meta-analyses and other information in Marsh et al. (2019) and Vincent et al. (2019) raise serious questions about the accuracy of USEPA’s characterization of the overall epidemiological evidence for EtO-induced lymphohematopoietic cancer and breast cancer as strong”. However, it should be noted that these references appear to present a reanalysis of preexisting data rather than adding materially to the substantive database in the period since the U.S. EPA (2016) assessment was completed.

Both Marsh et al. (2019) and Vincent et al. (2019) indicate that conclusions on breast cancer risk are based on 5 effect estimates (SMRs) almost all in sterilization facilities. Based on the content of Table 2 in Marsh et al., the only statistically significant increase (SMR = 2.55) was reported in Norman et al. (1995) and while the confidence interval for this study does not include 1, the range of SMRs is rather broad (1.31-4.98), compared to the other studies. While measures of quality in most studies were ranked as moderate or high by Vincent et al. (2019), these authors considered those for measures of exposure, outcome assessment and analysis in Norman et al. (1995) as low.

Based, then, on the relevant epidemiological data, there appears to be no or at best, limited evidence of consistency of an association between exposure to ethylene oxide and breast cancer mortality.

However, in my view, all of the discussion (such as that in Section 3.2) on the lack of concordance between tumor sites in animals and humans based on empirical data is not relevant to the assessment of the biological plausibility of the associations observed in epidemiological studies, though it’s presented in the TCEQ assessment as a basis to detract from support in this context. Reference here seems to be based on erroneous premise that empirical associations on site concordance are additionally informative in this context, where we have information on mode of action.

Site concordance between humans and animals is entirely a function of the relevant mode of action by which tumors are induced. Moreover, for chemicals assumed to act via a mutagenic mode of action – i.e., where mutation is an important early and influential key event, tumors are anticipated to occur at multiple sites (as noted by the U.S. EPA in their mode of action analysis cited in Section 3.3, Page 20 – lack of specificity in relation to site).

Based on the assumed mutagenic mode of action for ethylene oxide, tumors would be expected at multiple (though not necessarily concordant) sites in both animals and humans, with variation being a function largely of toxicokinetic and toxicodynamic differences. The observation of tumors at multiple (though sometimes discordant) sites in animal studies supports, then, the hypothesized mode of action of EtO and the biological plausibility of different tumors observed in epidemiological studies, rather than detracting from it.

I’d suggest, on this basis, then, that all reference to generic observations on species concordance of tumors based on empirical data within the document (Section 3), the responses to comments (e.g., page 23, para. 2) and the Summary be deleted. None of this contributes materially to the TCEQ rationale.

I suggest also, that the rationale for consideration of the evidence for causality of an association between exposure to EtO and breast cancer in epidemiological studies be strengthened by more systematic consideration of the extent to which available data fulfill widely accepted Bradford/Hill considerations, taking into account the observation of a range of tumors in animals, consistent with the hypothesized mode of action.

As per discussion above, the Bradford/Hill consideration of “consistency” of the association across studies between EtO and breast cancer appears not to have been met (though would be better informed based on consideration of the power of the studies to detect an effect), nor does that for “strength” of the association. (Authors from EPA have indicated that an increased risk of less than 50% (RR=1.0–1.5) or a decreased risk of less than 30% (RR=0.7–1.0) is considered by many epidemiologists to be either a weak association or no association

https://www.who.int/water_sanitation_health/dwg/nutrientschap9.pdf.)

The consideration of specificity of the tumors is not relevant for substances assumed to be mutagenic carcinogens. Those for temporality and dose-response across studies would be best considered though the concordance analysis mentioned below and in the responses to Questions 1 and 2 though there is evidence of dose-response within studies and that for biological plausibility is met, with a range of tumors being observed in animal and epidemiological studies consistent with a mutagenic mode of action. Consideration of coherence requires analysis of the concordance of dose-response relationships across animal and human studies, including incidence data (as per mention above and in responses to Questions 1 and 2).

Expert 5

It is appropriate that TCEQ did consider breast cancer as a basis for deriving regulations and devoted an entire appendix to that issue. It is not clear that breast cancer is indeed a hazard associated with EtO exposure, based on the recent studies that have been cited. And, I agree with TCEQ’s determination, based on the lack of site concordance reported by IARC, that animal observations of mammary tumors do not imply breast cancer risks will be exposure-related in humans. I am satisfied with the determination that decisions should be based rather on lymphohematopoietic cancers. A major contributor to my attitude is the expectation that appropriate handling of lymphohematopoietic cancer should also lead to decisions that are protective against other cancer risks, including breast cancer risk.

Expert 6

I would not characterize the evidence for EtO increasing the risk of breast cancer in occupational concentrations as convincing, but suggestive. If EtO does increase the risk of breast cancer at occupational concentrations, it probably also does at environmental concentrations, although not necessarily at a high enough level that can be detected, by virtue of evidence that it can induce a wide range of genetic damage, which increases the likelihood that it has a linear dose response. Also, see my response to question 13.

Charge Question 5

While it is in the interest of public health to protect against cancer *incidence*, available epidemiological studies often only provide cancer *mortality* data for dose-response modeling. What is your opinion on the accuracy of using a dose-response model based on cancer mortality data (e.g., lymphoid cancer mortality) to predict cancer incidence (e.g., lymphoid cancer incidence)?

Expert 1

Mortality data are sufficient indicators for cancer risks for malignancies that historically have had high case fatality ratios, such as acute myeloid leukemia. For breast cancer, which has a relatively low case fatality ratio, especially given the increasing practice of breast cancer screening techniques, mortality often gives an incomplete indication of population risks compared with incidence rates.

Expert 2

There is ample precedent for using cancer mortality data in the absence of appropriate data on the incidence of the cancer of interest. There are corrections to the estimates that can be made given knowledge of the mortality vs. the incidence of the particular cancer in the population being evaluated. For example, in the US EPA regulations on arsenic in drinking water, there was discussion of the observation that data modelled were for bladder cancer in a Taiwanese population. At the time of the study, bladder cancer was generally fatal in this population. This was not the case in the U.S. population to which the assessment was applied. In this instance the majority of the discussion of mortality vs. incidence of bladder cancer in the U.S. was in the context of cost / benefit of the proposed rules.

I do not have sufficient expertise on the treatability and mortality rates of the EtO associated lymphoid cancers to comment further.

Expert 3

As noted, it is preferable to consider cancer incidence as opposed to cancer mortality data for dose-response modeling, but making use of the best available evidence (mortality data in this instance) is necessary. The loss of information in going from incidence to mortality data results from:

1) Variable time course from case identification (incidence) to death, with the potential for years of delay between the events. This may result in exposure changes due to ceasing work, for example, at the time of diagnosis such that exposure in the period between diagnosis and death may be affected. Lagging exposure for some time before the event mitigates this problem, but does not take the concern away entirely since the interval between diagnosis and death is variable.

2) Influences on survival following diagnosis will affect the measure of association between EtO exposure and cancer mortality. At the extreme, with long-term survival, the cancer will not be identified at all but even when the disease is ultimately fatal, characteristics that predict lower mortality will distort the apparent etiologic impact of exposure. If younger people, for example, have more favorable cancer survival following diagnosis, they will appear to be at lower risk of cancer than is actually the case. Nonetheless, there are relatively few predictors of breast cancer survival that are likely to markedly distort the exposure-cancer relationship and a number of those such as age, ethnicity, etc. are controlled through statistical analyses.

3) Statistical power is reduced by the smaller number of events when studying cancer mortality as opposed to cancer incidence. This does limit the precision of estimates and is yet another reason (see below) to challenge the single-minded focus on statistical significance as the basis for reaching a

judgment. Given that the epidemiologic studies are all limited by imprecision, this is a real loss of information.

4) Given an interest in cancer incidence, non-fatal cases reflect a form of false negative assignments. Every individual who has the disease and does not die from it constitutes an error in this sense and depending on the pattern of those errors, they may tend towards spuriously elevated, depressed, or null associations or have no impact at all. But conceptually at least, those are errors.

Taking all of these issues into account, there is a clearly a loss of information in having to rely on cancer mortality data, but it is a reasonable proxy for cancer incidence overall and appropriate to use. As with all the other considerations bearing on the ultimate risk assessment, imperfect information is being assembled to make the best possible judgment. Note that this uncertainty does not lead to a predictable direction of bias, i.e., comparing the observed measures of association for mortality with the (unknown) measures of association for incidence, they may diverge to some extent but there is no straightforward way to assess whether the association for mortality would be over- or under-stating the association for incidence. Insofar as using the upper 95% confidence limit does in fact incorporate uncertainty, it does help to account for the extrapolation from mortality to incidence data but there's no logical basis for choosing this particular statistic to reflect that uncertainty or assuming that the direction of concern is upward when it's just as likely to be downward.

Expert 4

Clearly, wherever possible, cancer incidence data for specific tumors are preferred for dose-response modelling. Where only cancer mortality data are available, there should be some attempt to “ground truth” or “adjust” the estimates based on additional information concerning age adjusted ratios of incidence to mortality to estimate the extent to which risks are underestimated (i.e., to at least crudely quantitate uncertainties). Where this is not possible, as a minimum, this source of uncertainty needs to be addressed qualitatively. Where both are available for a range of different studies, it's often also helpful to analyze the relationships between the two, as a measure of consistency of the empirical associations with expected patterns. Selection of the upper 95% confidence of the slope doesn't address directly the variation between incidence and mortality which is best informed, to the extent that it can be, based on data on the variation between the two for the endpoint of interest. If there is a policy choice to select the upper 95% confidence interval to incorporate additional conservatism to address the uncertainty in relying on mortality data, this should be stated.

Expert 5

To me, it is not a matter of “accuracy” nor should this concern be couched in those terms. If the endpoint being modeled is mortality, then the modeling is only accurate (if even then) for mortality risks. That is, if the data on hand (from the studies of the cohorts suitable for dose-response analysis) relate exposure to cancer mortality, there is nothing one can do, to get morbidity-related risk estimates, that does not involve additional assumptions.

The question really should be, if we derive risk estimates relevant to mortality, how does this relate to risks of morbidity (incidence), and should extra “safety factors” be applied to achieve the desired level of protectiveness for morbidity. A simple step in the direction of that protectiveness would be to use lifetable methods in which the background rates are for incidence rather than mortality. It would need to be noted however, and as carefully documented and justified as possible, that the exposure-related increase (relative risk) used to derive extra risk estimates is based on relative risks associated with the disease's mortality. I.e., the assumption would be that the exposure-related effect on incidence would

be the same as the effect on mortality. I personally have no idea how good an assumption that would be, and it might depend on how fatal the cancer in question is and the time-course from incidence to mortality. As a technical note, if doing such a lifetable analysis, the survival rates from one age group to the next would have to be adjusted (downwards) to reflect the additional loss (from follow-up) of those individuals who did have an incident case of the endpoint in question, while being adjusted (upwards) to take out the loss from follow-up of mortality from the endpoint in question (which would already be accounted for because such loss would have been accounted for with the addition of the incident cases).

Expert 6

I am not an expert in the relationships of incidence to mortality for specific tumors. Obviously, the mortality risk should be smaller than the incidence risk. The difference should be particularly large for cancers that have a fairly large cure rate. Mortality statistics will be affected by improvements in the treatment of the cancer, whereas incidence statistics will not be.

Charge Question 6

The TCEQ's DSD discusses a problem with key USEPA AIC and p-value calculations used as criteria in determining model fit, and the TCEQ recalculated these values (DSD Section 3.4.1.3 and Appendix 4). Please explain what you think the appropriate approach should be for accounting for the number of estimated parameters in the modeling and the associated calculation of the AIC and p-values. Given that appropriate AIC and p-values are available for models fit to *individual* data, what role should visual fit to *categorical* estimates play in model selection (Response to University of California at San Francisco, Comment 6 in Response to Public Comments Document)?

Expert 1

Outside my area of expertise

Expert 2

I found the DSD section on visual fit to be difficult to pick apart. Nevertheless, the TCEQ arguments seemed reasonably convincing.

See specific comments below.

"Appendix 4 ~~Corrected~~ p-Values and Akaike Information Criterion (AIC) for the Two-Piece Spline Model and Other Models"

p. 136. - Suggest using "recalculated" or some other non-judgmental adjective throughout the document, rather than implying that TCEQ knows the true values.

p. 139 par 1 - "However, as use of an overall supra-linear model (i.e., the steep lower-dose slope) is not scientifically justified (see Section 3.4.1.4.1), the two-piece spline models are not considered for adoption; nor are other models that have an inherently supra-linear dose-response over the exposure range"

The DSD needs to make explicit here that these statements are conclusions of TCEQ, rather than received wisdom.

Expert 3

This question is outside the scope of my expertise.

Expert 4

Criteria for model fit are secondary to the consideration of the extent of the mechanistic base for modelling. This is recognized formally in weight of evidence considerations for integrating constructs such as AOP/MOA (across different levels of biological organization and study types such as in vitro, in vivo animal, clinical and epidemiological). For MOA/AOP, in application of modified Bradford Hill considerations for the extent of the evidence, biological plausibility is weighted most heavily (e.g., Meek, Current Opinion in Toxicology 2017, 3: 80–86, references cited in response to Question 2), followed by essentiality of key events (i.e., intervention studies to prevent or modify key events). Empirical support is ranked below these other two considerations (since association does not imply causation).

I also question whether model fit in the region of observation ever meaningfully informs extrapolation to the low dose region, without incorporation of quantitative biological information (e.g., biologically based dose response or case specific models involving quantitative modelling of key event relationships).

And while this is not my specific area of expertise, as I understand it, AIC penalizes models for adding parameters which do not significantly improve model fit. Authors from EPA indicate for BMD modelling, to use the smallest AIC, even when the differences are very small, as a basis to prevent users from choosing models “based on subjective and inconsistent criteria”. It further indicates that determinations of the most appropriate models are the global goodness-of-fit value (p-value), a measurement of local fit (χ^2 scaled residual values for each individual dose group), and a visual inspection of the model fit, though criteria for weighting are not addressed, to my knowledge:

https://www.researchgate.net/publication/47633481_Introduction_to_benchmark_dose_methods_and_US_EPA's_Benchmark_Dose_Software_BMDS_version_211 (Davis et al., 2010)

Based on program experience in the application of epidemiological data in characterizing dose-response for hazard characterization for a number of data rich environmental contaminants, individual data are not only preferred, but pretty much essential. It's difficult to envisage cases where categorical data would ever be preferred.

And as per comments in response to Question 10, in my view, it is the content of the subsections of Section 3.4. (principally 3.4.1.4) that describes TCEQ's approach and associated rationale that is the critical focus, here. It's unfortunate that it tends to get lost in secondary presentation to the extensive (often repetitive) text countering EPA's assessment.

In relation to the content of Section 3.4.1.4 (Selection of the Extrapolation Model), consistent with comments in response to other questions here, endogenous production and over prediction of EPA's model for the NIOSH cohort doesn't add weight in my view, to TCEQ's rationale for model selection (see, for example, the response to Question 10). Rather, in the absence of adequate mechanistic data to support otherwise, for extrapolation, a linear model has been adopted (presumably, consistent with TCEQ's documented science policy). Also, as per responses to other questions, it would be helpful to present by reference to previous examples, what TCEQ considers to represent “adequate” mechanistic data.

Inclusion of several approaches to “ground truthing” or “reality checking” of TCEQ’s estimates more reasonably informs, in my view. To the extent possible, this should draw upon the totality of the epidemiological evidence and toxicological data (see response to Question 8).

Expert 5

I do believe that TCEQ has identified a real problem with the USEPA AIC and p-value calculations. The explanation of the issue and the resolution supplied in DSD seems appropriate. That is, I agree with TCEQ that the knot parameter in the spline models should be considered in the count of the parameters, that the AICs reported by USEPA for those models are too low by a value of 2, and that the p-values should be computed using an approximation to a chi-square with 3 degrees of freedom.

The role of visual fit is not always an easy one, especially when the data are complex, as in the present instance. I have always tended to rely more on statistical diagnostics rather than visual fits, which are subject to numerous possible manipulations (choices of scale, resolution, etc.). I believe that the DSD actually, unintentionally, provides an example of the obfuscation that can be associated with visual evaluations. In Appendix 5, Figures 19 – 22 have 3 different “scales” for the y-axis, including what appears to be some non-standard multiplicative scale, to say nothing about the 2 different x-axis ranges.

There is an important corollary to all of these comments. The revised (in my opinion, “corrected”) AIC and p-values for the USEPA spline model put that model pretty much on the same basis as the proposed TCEQ model. That is, the AICs for the two models are pretty much the same; the difference is less than 2 which is a value often cited for distinguishing among different models. I.e., the rule of thumb is that an AIC difference of less than 2 suggests no reason to favor one model over another. That being the case, the question I would pose to TCEQ is: why would I prefer its model over the USEPA model when the latter is clearly more health protective? Of course, issues of prediction of observed responses (goodness-of-fit), simplicity (parsimony), and alignment with biological processes (MOA determinations) are important and should help prioritization in model selection. But in this case, none of those factors militates against the USEPA modeling approach. Thus, health protectiveness becomes a salient factor. And, to me, this is the bottom line for my overall opinion – for all of the denigration and adversarial tone in the DSD, I still cannot get past the fact that TCEQ has failed to show that their approach is better, methodologically, and that it appears driven solely by the desire to come up with a less protective result.

Expert 6

I consider that the location of the spline should be considered a parameter when evaluating fits of spline models, as long as the data were used in determining the knot, as it apparently was in EPA’s model. I believe also that the lag should also be considered a parameter when the data are used to determine its value. But, in general, I consider the AIC in such complex models to be essentially only a rough guide to evaluating fit. Therefore, I think TCEQ’s conclusion that the *“lower AIC means that TCEQ’s selected model is a statistically superior model fit than USEPA’s selected model when taking into account model complexity”* is an overstatement. Comparing a model with an AIC = 464.5 to one with an AIC = 264.4, you can only conclude with confidence that the two models fit about equally well. Additionally, the overall fit is not of major importance – the fit at small doses is much more important when the object of the fitting is to estimate the risk at very small doses.

I believe visual fits to categorical data, if presented in an equitable manner can be a useful tool for understanding the reasonableness of various fits. However, I would give greater reliance to formal goodness of fit tests.

However, I have concerns about TCEQ's Appendix 5 Visual Fit to the Underlying NIOSH Data. EPA's categorical RRs (shown as five red dots in the figures) were estimated using the same approach as the TCEQ Cox model, as I understand what was done, except a different parameter was estimated for each exposure category and the exposures used for each subject was the same representative value for the exposures for the category in which he/she belonged. Consequently, EPA's categorical RRs are relative to the unexposed category, as are TCEQ's Cox regression and EPA's spline model. Therefore, all three of these should be comparable when graphed. I don't understand how the nonparametric rate ratios for individual cases were estimated and did not find a description of their estimation. With the possible exception of these points, it should be a fair visual comparison to see EPA's categorical regression RRs, TCEQ's Cox regression RRs and EPA's spline RRs all on the same graph, since they are all relative to the same background response in unexposed subjects. So, it seems to me that TCEQ's argument about these different estimates not being comparable because they are based on different baseline risks doesn't hold water. I also think that TCEQ's objections to EPA's comparison of the spline model to the categorical data points (EPA Figure 4-3) due to the categorical data not being the data used for fitting the models are overblown, because the categorical data points are a summary of the underlying data. I don't understand why in Figure 22 the Cox regression was adjusted (by multiplying by RRo), thereby changing the estimate at zero dose to RRo, while the EPA spline model and EPA's categorical RRs were not adjusted and equaled 1.0 at zero dose. Based on the linear regression of the nonparametric rate ratios for individual cases (as stated earlier I don't know how these were calculated), apparently they were not relative to the unexposed category (because the RRo estimate is quite a bit greater than 1.0) as were EPA's categorical RRs, the Cox regression, and the EPA spline model. Therefore, I don't see the logic for adjusting only the Cox regression in Figure 22.

Charge Question 7

Please comment on the biological and mechanistic support for and against use of an overall supralinear model to estimate risk of lymphoid cancer from exposure to EtO at occupational levels and at environmentally-relevant concentrations.

Expert 1

Outside my area of expertise

Expert 2

Neither TCEQ nor US EPA provide biological or mechanistic justification for use of a supralinear model for extrapolation in the low dose region of the EtO dose response curve for cancer. Given the discussion of model fit provided in the DSD it makes no sense to apply a supralinear model. If conservatism is a goal of the TCEQ URF, then a linear model can reasonably be used. If biological relevance is the goal, less-than-linear approaches could be considered. Throughout the DSD (and at least in some places in US EPA 2016), there are statements supporting observations of non-linearity at low dose.

Expert 3

While this is outside the range of my technical expertise, from my review it seems that the burden of proof for invoking a supralinear model to estimate risk is on those who postulate it rather than being the default assumption that must be disproven with evidence. The document makes a convincing case that there is not a compelling justification for invoking a supralinear model based on mechanistic or epidemiologic evidence.

The judgment regarding the presumed shape of the dose-response curve in the range of concern is not going to be determined empirically but rather inferred based on available lines of evidence. Triangulation is necessary taking into account the presumed mode of action, empirical evidence on risk of cancer in relation to exposure that can be quantified, and then circling back to how the limits that are being considered line up against other considerations such as endogenous levels and predicted population impact.

Expert 4

In my view, there is limited mechanistic data to justify deviation from linear extrapolation in this case and as I understand it, EPA has not provided a rationale for same. With the exception of the possible impact of additional analyses of the dose-response relationships for early key events in the hypothesized mutagenic mode of action, empirical data also do not inform this discussion (given the high exposures of workers in observational epidemiological studies and the high doses to which animals have been exposed in bioassays).

If the analysis of concordance of dose-response relationships across several levels of biological organization and data sources (e.g., in vitro, in vivo animal and human), does not meaningfully additionally inform extrapolation to the low dose region, the decision concerning selection of the relevant model is a function largely of policy, rather than scientific judgment. (As per the response to Question 10, the arguments on endogenous formation are not particularly convincing, in determining the most appropriate model for extrapolation).

It appears that EPA acknowledges that their decision to adopt a supra-linear model constituted a policy choice for additional conservatism, based on the observation on page 27 (first para.) that *“It is also critical to note that USEPA acknowledges the lack of mechanistic data to support the biological plausibility of an overall supra-linear dose-response, stating “the EPA is not aware of a mechanistic explanation” in response to questions from the USEPA SAB (p. I-29 of US EPA 2016)”*.

In this context, TCEQ indicates (page 25, first full para.) *“The TCEQ guidelines (2015) go on to state... “Using supra-linear exposure-response models can only be justified if there is sufficient biological or mechanistic data to support their application.” Another way to state this more specifically might be [added]... “Using the initial steep slope starting at zero dose in supra-linear exposure-response models can only be justified if there is sufficient biological or mechanistic data to support their application.”*

What might be helpful, then, as a basis to increase transparency in the basis for selection of a model for extrapolation is comparison of the considerations being made by TCEQ for EtO in this case versus those where they have selected a supra-linear model for extrapolation (if there are any such examples). It is experience of this nature on which criteria or considerations for future cases (i.e., guidance) are normally based.

This, in my view, would provide a stronger rationale than is presented currently and appropriately transparently acknowledge science policy choices and precedent of accumulated experience in the absence of relevant observational data. For example, in para. 2, page 25, it is indicated *“In the present case, the TCEQ finds insufficient data to justify the supra-linear modeling approach (i.e., use of the steep lower-dose slope starting at zero dose from the linear two-piece spline model) ultimately adopted by USEPA (2016)”*. What does or has TCEQ considered as “sufficient” data to justify such an approach (What are the criteria or considerations?)

Expert 5

My response to this question is contained in the lengthier response to question 3. In a nutshell, TCEQ has misconstrued and misrepresented the whole notion of supra-linearity. Even though I am not a toxicologist or biologist per se, I have discussed above some potential reasons why the slope of the relative risk of response might decrease as dose increases, even for a direct-acting carcinogen like EtO. And I noted that the biologically based Michaelis-Menten equation has exactly the shape that TCEQ finds objectionable, i.e., one that is steeper (though not infinitely steep) at lower doses than at higher doses.

My overall conclusion is that in the DSD I see no well-developed biological or mechanistic objections to using a piece-wise linear spline method, even if the resulting shape is steeper at lower doses. The DSD is not convincing with respect to its objections to the modeling done by USEPA, at either occupational levels or at environmentally relevant concentrations.

Expert 6

The assumption by TCEQ that the EPA risk assessment model is “supra-linear” is the TCEQ’s chief reason for rejecting the EPA assessment. However, this reflects a misunderstanding of the term supra-linear and is a mistake. The term “supra-linear” has two different meanings in the scientific literature. One of these meanings is (or appears to be for I haven’t located a formal definition) simply that a straight line drawn from a point on the dose-response to the response at zero dose lies below the dose response in this dose region. It appears that this is the definition (“Def.1”) that TCEQ has in mind, judging from their Figure 1 that illustrates a supra-linear dose-response. It appears that this is also the definition of supra-linearity that EPA had in mind in their EtO risk assessment.

By contrast Crump (2011) gave formal definitions for four low-dose behaviors: low-dose supra-linear, low-dose linear, low dose sub-linear and threshold: A supra-linear low-dose response is one that has an infinite slope at zero dose (i.e., an infinite first derivative at zero dose) (Def. 2). According to these definitions, a low-dose linear dose response is one that has a finite positive slope at zero dose (i.e., a finite positive first derivative at zero dose); a threshold dose-response is one in which there is some positive dose such that the response at that dose and all smaller doses equal the background response; a low-dose sub-linear dose response is one that is not a threshold but the slope at zero dose is zero (e.g., $\sim \text{dose}^2$). Note that these definitions categorize dose-responses according to their characteristic in the neighborhood of zero dose (i.e., in the low dose region) which is the region of interest in low-dose risk assessment. Note also that they are independent of any data used in deriving the dose-responses.

By comparison, the EPA Cancer guidelines (US EPA 2005) defines low-dose linearity in the same way as above: as a dose-response “whose slope is greater than zero at a dose of zero.” Rather than separately defining low-dose sub-linear and threshold, US EPA (2005) defines “low-dose non-linear” as a dose-response whose “whose slope is zero at a dose of zero.” Note that this includes both low-dose sub-linear and threshold dose responses as defined above but does not include supra-linear dose responses as defined by either definition. The EPA guidelines do not discuss or define supra-linearity.

There are papers in the literature that state, or at least imply, that supra-linear dose responses according to Def. 2 are biologically implausible and therefore should not be used for low-dose risk assessment (e.g., Crump 1984, 1995, 2011). However, dose-responses corresponding to Def.1 are not necessarily biologically implausible. In fact, with Def. 1, even the one-hit model, $P(d) = 1 - \exp(-a \cdot b \cdot d)$, the prototypical low-dose linear model, would be labeled “supra-linear.” In particular, the EPA EtO model is exactly linear at doses below the knot, and therefore would be termed a low-dose linear model by any

reasonable definition of linearity including the definition in Crump (2011), and therefore is not biologically implausible due to its low-dose properties.

TCEQ apparently took the statements in the literature that a supra-linear model is biologically implausible (according to Def. 2) and inappropriately applied them to the EPA's dose response, which is not supra-linear according to Def. 2. Moreover, the EPA spline model is not even supra-linear according to Def. 1. Below the knot, the dose response of the EPA spline model is a perfectly straight line. That must be "linear" by any definition. TCEQ is apparently confusing shape (e.g., linear versus supra-linear) with potency. The DSD refers repeatedly to the "steep lower-dose component" of the EPA model and identifies this as evidence of supra-linearity. Steepness relates to potency, not to shape. The EPA spline model is not supra-linear according to either definition of supra-linearity (Def. 1 or Def. 2). Apparently, the DSD is basing its decision that the EPA model is supra-linear partially on the visual appearance (e.g., TCEQ Figure 14). However, that visual appearance depends on the exposure scale. On a different scale the steep visual increase in the low dose range would disappear.

There are places in the published literature where supra-linear dose-responses are stated to be not biologically plausible, and in every case (of which I am aware) these comments are referring to Def. 2 supra-linearity (e.g., Crump 1984). The discussion in the DSD around Crump and Allen (1985) misunderstands what this paper means by a linear dose response. By "linear" the paper means linear according to the definitions given above – a dose-response that has a positive, non-infinite slope at zero dose. In saying that linear models are "considered conservative in the sense that other biologically plausible dose-response models would generally imply lower risks," by "biologically plausible" Crump and Allen are ruling out models with an infinite slope at zero dose (Def. 2 non-linearity) but including curve shapes that have a finite slope as zero dose (low-dose linear, low-dose sub-linear and threshold). This is consistent with the definitions in Crump (2011) given above.

Therefore, statement like that on page 5 of the DSD "Supra-linear models are generally not biologically plausible and tend to grossly overestimate low-dose risks" (unreferenced) are misleading because such statements in the literature presumably refer to models that are supra-linear according to Def. 2, but TCEQ is applying them to models that are supra-linear according to Def. 1. At any rate such discussions are of no consequence because the EPA model doesn't satisfy either definition of supra-linearity (Def. 1 or Def. 2) and consequently the EPA model cannot be considered biologically implausible because of supra-linearity.

My short answer to Question 7 is, therefore, properly understood, both the EPA spline dose-response model and the TCEQ Cox model are low-dose linear rather than supra-linear (by either definition of supra-linearity) and consequently the EPA model is not inappropriate or non-biological on that account.

TCEQ missed the point of USEPA's statement that "EPA considers it highly plausible that the dose-response relationship over the endogenous range is sublinear" (e.g., DSD, page 4-95). But "sublinear" refers here to the endogenous range, not to an actual exogeneous dose-response that can be viewed and studied. EPA explains this on the basis that baseline levels of DNA repair enzymes and other protective systems evolved to deal with endogenous DNA damage could work more effectively for lower levels of endogenous adducts. However, as explained in Crump et al. (2014), a sublinear dose-response in the endogenous range leads naturally to a dose-response in the endogenous range with a positive slope at zero (exogeneous) dose (i.e., a low-dose linear dose-response). The figure in Crump et al. (2014) illustrates this phenomenon. The important point is that the fact that EPA claims that the dose-response relationship over the endogenous range is likely sub-linear does not imply that the observable dose-response in the exogenous range is sub-linear. In fact, it suggests that it is linear.

Put another way, if endogenous exposure is responsible for any of the background tumors (whether by a linear or sub-linear dose response in endogenous exposures) then the slope of the dose-response at the interface between endogenous and exogenous exposures (i.e., at zero exogenous exposure) will be positive. Therefore, the dose-response as a function of exogenous exposure will be low-dose linear (Crump et al. 1976, or, for a more recent discussion, Crump 2017).

The DSD, in discussing endogenous and exogenous exposures, states that *“USEPA (2016) applied remarkably steep supra-linear model low-dose slopes for lymphoid and breast cancer (see Figures 4-9 and 4-10 of USEPA 2016) in the very region where sub-linearity is expected (i.e., \leq the normal endogenous background range).”* I don’t understand this statement. EPA argues that sub-linearity is expected in response to endogenous exposures. EPA did not model the response to endogenous exposures, but instead modeled responses to exogenous exposures. EPA did not claim that the response to exogenous doses should be sub-linear, in fact, as suggested in the previous paragraph, they argued that the response to exogenous should be linear (linear, as defined in Crump 2011).

Likewise, it seems to me that Figure 7 in the DSD is misleading, or perhaps I miss the point. It shows exposures corresponding to the EPA target risk ranges, along with higher exposures corresponding to endogenous levels. The point seems to be that exposures corresponding to the EPA target risk ranges, are smaller than the exposures corresponding to endogenous levels. But the exogenous exposures are added to the endogenous exposures. Therefore, I see no basis for a direct comparison since the exogenous exposures are added to endogenous exposures and never occur alone.

Charge Question 8

As summarized in DSD Section 3.4.1.4.2, the TCEQ used MOA, model predictiveness reality checks (both for the NIOSH cohort and the general population), biological plausibility, and statistical model fit criteria for model selection. Have these considerations been clearly described and are they scientifically appropriate given the available data?

Expert 1

Outside my area of expertise

Expert 2

The checks on model predictiveness and so-called reality checks are reasonably done, but in several places the stated conclusions are undermined by what I am calling judgmental language and over-hyping. Please see marginal notes on text of the marked-up copy of the DSD.

Expert 3

The use of these “reality checks” is well-justified and clearly presented. In general, modeling of this sort can generate implausible results given the many untestable assumptions underlying it, and seeking multiple approaches to judging whether or not they are sensible is critically important. Relating predicted effects of EtO on disease occurrence to the predicted impact of endogenous exposure levels, occupational exposures, and community exposures is very helpful to judging the general reasonableness of the model. While such exercises cannot confirm that the model is correct in an absolute sense, it can raise a red flag to indicate when something is clearly erroneous. Interpreting the degree of correspondence between the model predictions and what is actually occurring should take into account statistical power to identify the model-predicted effects.

Expert 4

I appreciated TCEQ's attempts to "reality check", "benchmark" or "ground truth" their estimates of risk, based on the selected model. However, I wondered if it wouldn't be possible to include additional analysis of dose-response relationships across the toxicological and epidemiological database, taking into account the hypothesized mode of action. (This follows from suggestion to more robustly consider concordance of dose-response relationships for key events in the hypothesized mode of action in responses to other questions).

How consistent are the estimated risks with what has been observed in epidemiological studies of cancer mortality other than the NIOSH cohort? Would it not also be possible to consider the consistency of the predicted risks with observational data in humans on earlier effects such as cytogenetic changes?

In addition, while the database was more limited at that time, and there are necessarily limitations of such analyses, the consistency of the estimates of cancer risk from studies in EtO-exposed animals were also compared to those for hematological cancers in the NIOSH cohort in the Health Canada (2001) assessment:

https://www.canada.ca/content/dam/hc-sc/migration/hc-sc/ewh-semt/alt_formats/hecs-sesc/pdf/pubs/contaminants/psl2-lsp2/ethylene_oxide/oxyde_ethylene_oxide-eng.pdf

Expert 5

My take on these issues is summarized as follows:

a. Nothing related to MOA or biological plausibility has convinced me that the USEPA method is inherently flawed. I believe that the dose-response should be essentially linear at low enough doses. The USEPA model is low-dose linear. The question that I have raised in this regard is if the slope of that linear portion of the curve can be estimated from (set equal to) the slope derived by considering the occupational exposure levels that are much greater than those that will correspond to a truly low-dose region. The DSD is not convincing in that regard, as discussed above. TCEQ missed the opportunity to address this issue by excluding from consideration (by the very nature of its literature search) data relevant to low-dose behavior that might have shed light on, and helped estimate, any difference (presumably reduction) in slope at exposure levels below those observed in the epidemiological datasets.

b. With respect to statistical fit criteria, all that the DSD shows is that the USEPA model and the proposed TCEQ model are roughly the same. As stated elsewhere, why would I pick the less protective option (TCEQ's model) when the statistical fit is essentially the same?

c. The so-called model predictiveness reality checks are not convincing. In response to question 3 above, I laid out a set of counter-arguments that suggests that the "endogenous EtO" reality check does not argue strongly against USEPA's URF. The reality checks related to predicting the number of responders predicted by the models (Appendix 2 of the DSD) are problematic for me, as I will attempt to explain here.

First, there is this statement (p. 109 in Appendix 2): "There is no fairer evaluation of the predictiveness of a given model assessment than direct numerical comparisons of the specific model's predictions to the reality of the dose-response data." Not true. A fairer evaluation would be to compare the predictions to data that were not used to estimate the parameters of the model. So, right off, I am wary of this "reality check."

Second, the methods used to get the predicted numbers from the models are not given. If this is common knowledge, I apologize, but it is not apparent to me how the predicted numbers were computed. Moreover, given the above-noted similarity between the TCEQ- and USEPA-selected models in terms of fit statistics, it is not clear how they could differ so substantially with respect to predicted numbers. My suspicion of some calculation errors cannot be allayed without knowing how the numbers were computed.

Third, I do not understand how confidence bounds on the predicted numbers were derived. This concern arose when I noted that some of the predicted numbers were based on using the 95% upper bounds on the model parameter values, and yet those predictions are themselves presented with confidence bounds. How can there be an “additional” uncertainty about the predicted numbers that is seemingly added on top of the uncertainty in the model parameter values? If those bounds are based on the equations for the bounds on an SMR shown in Section A2.3 (p. 116), then I wonder if there has been some confusion about what the variable “E” is in those equations. If I understand correctly, E is the expected number from the “control” group (or background population). It is not the model-predicted number. My understanding is that in those equations, O and E would be fixed (model-independent), and that those calculations represent “sampling error” – equations to account for the finite sample of O observed cases. It does not appear that those equations should have E as a function of the exposure-related changes in RR.

TCEQ needs to clearly delineate and explain how these calculations were made and the rationale behind them. In conclusion, I find the arguments in the DSD to be neither clearly laid out nor scientifically appropriate.

Expert 6

My response is numbered corresponding to the numbers in DSD Section 3.4.1.4.2:

1. Section 3.4.1.4.2 cites the fact that the Cox regression model is linear, as appropriate for a direct-acting mutagenic carcinogen. I agree. However, the same can be said for the EPA model (see the response to the previous question).
2. I agree that accurate predictions of cancers over the entire dose range are a desirable feature of a model. However, the low-dose predictions of a model are much more important for cancer risk assessment than the prediction over the entire exposure range.
3. The prediction of the lymphoid cancer rate in smokers resulting from endogenous EtO appears to be very uncertain and not a strong basis for disregarding the EPA model. The Kirman and Hays (2017) reference contain two estimates of the endogenous level in smokers, the smaller of which is only marginally larger than the estimate for non-smokers. The TCEQ discussion assumes the larger estimate of the endogenous EtO level in smokers.
4. Biologically meaningful doses is an undefined term. It seems entirely possible that exposures could correspond to the very small risks of interest (10^{-6} to 10^{-4}) and still not be considered “biologically meaningful” (see response to question 10).
5. I agree that, based on the evidence presented in the DSD, the TCEQ Cox model appears to present a slightly lower AIC than the EPA spline model when considering the fit to the entire data set. However, the differences are very slight and the most one should conclude is that they provide roughly comparable fits (see response to question 6). However as stated elsewhere, fidelity with the low dose

data is a better criterion to judge the applicability of the resulting risk estimates corresponding to very low exposures.

As an addendum to my original comment I offer suggestions about defining what I mean by “low dose data” and how to determine them. By low dose data I am referring to occupational data in a dose range from zero dose up to some maximum dose in which the occupational data in this range are compatible with a linear dose response, but also contain enough data points so that the linear slope (or the unit risk) defined by the data is not overly uncertain. One possible way of defining this range is to first divide the dose range up into small increments (e.g., 20 increments). Then, starting with the lowest dose increment, sequentially add the data at the next lowest dose increment, and with each added dose increment conduct a goodness of fit test of whether the included data are consistent with a linear dose response. Continue adding increments until the included data are no longer compatible with a linear dose response according to the goodness of fit test. The added data minus the last increment added could perhaps be a reasonable set of “low dose data” for estimating the unit risk. Alternatively, one could consider starting with the complete data set and sequentially remove the remaining highest dose increment until the remaining data become consistent with a linear dose response. Another possibility for estimating a unit risk is to determine a benchmark dose (BMD) associated with a suitable benchmark response (BMR) and then simply draw a straight line from the point at the BMD to zero dose, as is frequently done with animal data. I am not sure that that any of these suggestions would work well for all circumstances, but this at least perhaps gives a better idea of what I meant by “low dose data” in my comments.

Charge Question 9

In DSD Sections 3.4.1.4, 3.4.1.5, and 3.4.1.6, the TCEQ describes their modeling choices and assumptions, and calculates an inhalation unit risk factor (URF), ultimately applying age-dependent adjustment factors (ADAFS) in DSD Section 3.4.2. Do you disagree with any of the modeling choices and assumptions or calculations made by TCEQ in the dose-response assessment? Please discuss any issues or concerns you have with the inhalation URF derivation.

Expert 1

Outside my area of expertise

Expert 2

I agree with the modeling choices and in general find description of rationales to be acceptable. I have noted areas that I found unclear, etc. in the marked up DSD.

It is a reasonable choice to apply the default US EPA ADAFS to the assessment – assuming that TCEQ has, in fact, demonstrated that a mutagenic MOA is appropriate for EtO carcinogenicity. However, the DSD should describe the scenarios under which an ADAF-adjusted URF would be applied. I am assuming that it would be applied when assessing the risks to the general population (not workers) exposed to EtO, but this is not made clear in the DSD.

I have included my comments on several appendices under this question (see below).

3.4.1.5 Relevant Cox Proportional Hazards Model Results

P 66, par 3 - See marginal note in the marked up DSD. I strongly suggest that this statement refer to TCEQ's own evaluation. If that is the same as that of US EPA, then add a statement to that effect. The main point is not that the lymphoid cancer data set best supports US EPA's classification, but rather that TCEQ finds that this data set provides the strongest WOE for its categorization of cancer.

3.4.1.6.1 Critical Cancer Endpoint

p. 75 last par - "As discussed in Section 3.2.2.1, the IARC (2019) unanimous consensus is that "At present, the state of the science does not support tumour site concordance as a general principle." Thus, current best available science indicates that animal data cannot generally be used to support specific sites of chemically-attributable carcinogenesis in humans. This is even more so the case when laboratory animal results are inconsistent; for example, when EtO induces mammary tumors in mice but not rats."

I suggest less reliance on the IARC statement. The inconsistent observation across species provides a more solid argument.

p.77 1st par - "the carcinogenic to humans classification is best supported by the lymphoid cancer data."

This reviewer is not convinced (at least not by the DSD as written) that "carcinogenic to humans" is the best supported descriptor.

3.4.1.6.2 URF and Air Concentrations at 1 in 100,000 Excess Risk

p. 77 2nd par - When determining the final EtO URF, the weighting of data from both cohorts (NIOSH and UCC) must be considered

Because? Is this TCEQ policy? Is there guidance? The presumed examples in the next sentences are not illuminating in their current form.

p.78 1st par - "As seen from Table 16, using person-years $\times 1/SE^2$ as a weighting factor results in the NIOSH (males only) cohort receiving ≥ 33 -fold greater weight than the UCC (males) cohort. Aside from consideration of cohort person-years or the number of cohort cancer mortalities observed, using $1/SE^2$ as a weighting factor produces qualitatively similar results, with the NIOSH (males only) cohort receiving >10 -times more weight than the UCC (males) cohort. Thus, based on the considerations inherent to the weighting factors applied, results suggest that for all practical purposes the URF (and corresponding 1 in 100,000 excess risk air concentration) should be based on the NIOSH cohort."

I found this unconvincing and abstruse, as per note on the marked-up text.

p. 78 par 2 - "Furthermore, as both a scientifically reasonable and conservative selection",

Is so-called "conservatism" a policy of TCEQ in selecting dose response values? I suggest that the more compelling argument is scientific reasonableness.

p. 78 par 3 - "... cohort would be somewhat higher at 5.2 ppb, but within a factor of 1.3."

Is this acceptable? Typical? What is the meaning of the 1.3?

3.4.2 Evaluating Susceptibility from Early-Life Exposures

p.79 last par - "USEPA (2016) indicates that there are no data on the relative susceptibility of children to EtO"

Are there any data from early life stage exposure in animals? Probably not, but this should be indicated.

*p.80 par 3 - "Note that this value would be insensitive to an additional 10-fold ADAF for in utero exposure during the third trimester (i.e., equation 5-17 of TCEQ 2015 would become $5.9E-06/URF = 5.9E-06/2.5E-06$ per ppb = 2.36 ppb). **The ADAF-adjusted URF is $4.1E-06$ per ppb or $2.3E-06$ per $\mu\text{g}/\text{m}^3$ [emphasis in original text]** (i.e., $2.5E-06$ per ppb \times 1.657 (based on equation 5-16 of TCEQ 2015) = $4.1E-06$ per ppb)."*

As noted in text comment above, the equation with definition of parameters should be given in the text, and then applied for EtO. I also recommend that the DSD explain the comment re insensitivity to additional 10-fold ADAF for *in utero*.

Appendix 2 - p. 117 par 3 - A2.3.1 "US Background Hazard Rates are Appropriate for Calculating the Expected Number of Lymphoid Cancer Deaths in the NIOSH Cohort due to Absence of a Healthy Worker Effect for Lymphoid Cancer Mortality"

Where are the data and / or citations that there is no healthy worker effect for lymphoid cancer mortality? Is it the Kirkeleit et al. (2013) paper mentioned later? This is one of the several situations in the DSD wherein a conclusion is stated as truth and then the arguments or support for it is given later.

A2.4 Calculating the Expected Number of Cause-Specific Deaths in a Cohort Assuming that the Death Rate in the Cohort Increases with Cumulative Exposure

P. 119 par 2 - "if the background hazard rate is assumed to be affected by exposure to a carcinogen via a multiplicative function"

What is the basis for this assumption?

p. 120 –

And then the Appendix just ends with no further explanation of the formulas or tables. Or conclusions to apply to the URF derivation or anything else.

Appendix 3

p.134 last par - "the OSHA PEL (1 ppm) is 222 times the air concentration corresponding to the 95th percentile of the normal endogenous background range"

What is the purpose of this comparison?

Appendix 4 ~~Corrected~~ p-Values and Akaike Information Criterion (AIC) for the Two-Piece Spline Model and Other Models

p. 136.

I suggest using "recalculated" or some other non-judgmental adjective throughout the document, rather than implying that TCEQ knows the true values.

p. 139 par 1 - "However, as use of an overall supra-linear model (i.e., the steep lower-dose slope) is not scientifically justified (see Section 3.4.1.4.1), the two-piece spline models are not

considered for adoption; nor are other models that have an inherently supra-linear dose-response over the exposure range”

The DSD should make explicit here that these statements are conclusions of TCEQ, rather than received wisdom.

Appendix 5. - A5.1 Non-parametric Rate Ratios are NOT the Observed Data

p. 142 last par. –

In the absence of some definition (or concept of the meaning of) rate ratios, I can't evaluate the validity of this section.

P. 143, par 2 - “Categorical rate ratios (RRs) should not be used for visually comparing models fit to individual data, particularly when appropriate statistical model fit criteria are available.”

I assume that this is a conclusion of TCEQ, based on either some assessment / calculation or some cited work. In the absence of some rationale, I can't judge the validity of this statement. In this section (and some others) it appears that the conclusion is presented before the arguments leading to that conclusion. This makes one skeptical.

Expert 3

Based on the information provided on modeling assumptions and calculations, which is fully documented and supported, the choices all seem to be reasonable ones. Given that they are ultimately judgments, some form of sensitivity analysis would be useful where the “right” decision is unclear and there are a range of reasonable possibilities to be considered.

Expert 4

Most content here (i.e., specific aspects of the modelling) does not lie within my area of expertise, with the exception of that for Section 3.4.1.6. While I agree that the extent of support for the causality of the association between EtO and breast cancer is limited and precludes it being considered a critical endpoint, I disagree with the rationale provided (points 1, 2 and 3 on page 76). (See response to Question 4).

In my view, the rationale is best predicated on consideration of the causality of the association, based on more systematic analysis of the widely accepted Bradford Hill considerations rather than misplaced emphasis (at least in my view) on generic hypothesis generating observational empirical associations of tumor concordance between animals and humans (particularly for EtO, where the evidence is relatively consistent with a mutagenic mode of action).

I wondered also if the analyses in Section 3.4.1.5.2.2 couldn't be used as a basis for comparison with the estimates for lymphoid cancers, with the objective to clarify that while breast cancer is not considered a critical endpoint, analyses indicate that the estimates for lymphoid cancer are also protective for this endpoint.

Expert 5

I do not disagree with the specific calculation of the URF or the application of the ADAF. And, it is not so much that I disagree with the modeling choices and assumptions, it is more that I do not think that the DSD has provided any reason not to rely on the more health-protective value derived by USEPA. There is

no basis (that I understand and can say was done correctly – see my response to the previous question) for down-weighting the more conservative estimates provided by USEPA.

As stated earlier, I believe one should consider a Bayesian approach that could integrate additional pieces of information via definition of priors for parameters in a model that had more “flexibility” in the low-dose region.

One might also consider a model-averaging approach, especially given the essentially identical statistical fits of the USEPA- and TCEQ-selected models.

Expert 6

The description of the ADAF calculation in Section 3.4.2 “Evaluating Susceptibility from Early-Life Exposures” is not clear. (E.g., the 6.0E-06 in the last line of page 79 of the DSD isn’t explained.) The description of applying ADAFs needs to be explained more clearly in the document before it can be evaluated.

Other concerns about TCEQ’s modeling are contained in responses to other questions.

Charge Question 10

Based on biomarker data, various sections of the DSD (e.g., Section 3.4.1.2.1, Section 3.4.1.4.2 number “4.”, second to the last paragraph of Section 3.4.1.6.2) discuss air concentrations corresponding to endogenous and background EtO levels and also compare these levels to acceptable air concentrations derived from URFs (either the TCEQ’s or EPA’s). Such a discussion is also included in the Response to Public Comments document (e.g. Response to Dr. Kyle Steenland, Comment 3). Please comment on whether the information and context provided by the discussion of endogenous/background EtO levels is clear and is scientifically appropriate.

Expert 1

Outside my area of expertise

Expert 2

I generally found the material to be appropriate. See some notes on marked-up copy of the DSD, as well as the comment below.

3.4.1.2.2 Key Epidemiological Data with Additional Context Using Endogenous Data and Model Predictions

P 32 par 3 - “If the underlying dose-response for EtO-induced cancer in humans were supra-linear with a steep low-dose slope beginning at zero dose, statistically significant increases in critical cancer endpoints would be expected beginning in the lower occupational exposure groups.”

I assume this would be true if the steep slope were applied to the whole dose response range. I would not expect this to be the case if there were a break in dose response, at which the higher exposures resulted in less steep slope. This dose response break can be seen with competing toxicity, competition for “activation” metabolism, etc.

Expert 3

Examining the URF in relation to endogenous and background EtO levels is quite helpful and seems to be scientifically appropriate. While I cannot comment on the technical accuracy of these calculations, the rationale is explained well and convincing. With regard to the EPA URF, juxtaposing it with endogenous levels of EtO and general community levels raises a real concern that something has gone wrong in their modeling and needs to be re-examined, which TCEQ has done. In trying to judge what levels are acceptable, it is not reasonable to assign endogenous levels as “unacceptable.” Likewise, a background level in communities that suggests EtO causes a vast proportion of lymphatic cancers is highly unlikely to be accurate and almost certain to be overstating its impact.

Expert 4

I think that the text in Section 3.4.1.2.1 (Endogenous Levels) goes beyond what is warranted based on available data. While the section nicely lays out the likely extent of minimal contribution of exogenous EtO to levels of circulating HEV adducts (stated to be for additional context, only), the “leap” to the likely biological impact (e.g., as in the following text in para. 3, pg. 29) is unjustified, in my view:

“More specifically, this suggests that inhalation exposure to sub-ppb EtO air concentrations, particularly concentrations in the range of parts per trillion (e.g., 0.1-10 ppt), is of little biological importance compared to normal endogenous background levels”.

This is also true of the following cited conclusion from Swenberg et al. (2008) (bottom of page 29, top of page 30):

“The biologic effects of de minimus exposures below endogenous amounts are lost in the noise of the background (e.g., carcinogenesis is driven by endogenous DNA damage when the dose-response for mutations due to external EtO exposure comes into the normal background frequency due to endogenous production)”.

While the extent of minimal contribution to exogenous load is relevant for consideration to set context, without additional mechanistic investigation of the relative extent to which exogenous and endogenous EtO contribute to circulating/ target tissue adduct levels (not so difficult to examine), it cannot be concluded that its biological impact is negligible or unimportant. (Given how much is made of this argumentation in relation to EtO, it’s surprising that the relevant data have not been generated or at least, I’m not aware of it). In my view, this line of reasoning rather discredits the value of other more justifiable components of TCEQ’s rationale.

I also question whether the content of Sections 3.4.1.2.1.1 “Reality Check Using Endogenous/Background Level Data” and 3.4.1.2.1.2 “Endogenous Conversion of Exogenous Ethylene to EtO: Potential Risk Implications based on USEPA (2016)” is really necessary. It seems a bit peculiar (at least to me) to be presenting the TCEQ analysis in a rather defensive context throughout the text investing rather more effort than is necessary in discrediting the EPA assessment. How much (if anything) does the current content of these two sections contribute in the context of defensibility of the TCEQ approach?

I’d submit that TCEQ has reasonably focused their assessment on areas of controversy, based on the most recent regulatory assessment (US EPA, 2016), an update of the literature review since that time and taking into account precedent and policy of TCEQ. It might be helpful to state this in the text in a formal problem formulation, which addresses context, objectives, resources, focus. etc. This is distinct from the much narrower context for “problem formulation” included for the literature review in

Appendix 1. Perhaps, the latter might more appropriately be entitled as Scope and Focus of the Literature Review. Experience and expertise of TCEQ to fulfill their mandate is well respected; countering the EPA assessment as the critical content of the assessment is unnecessary, in my view.

Rather, following delineation of the focus and scope of the assessment, it's sufficient to transparently document the basis and background for TCEQ's approach (which I believe has been largely achieved, with a few exceptions as noted in the comments, here). Often, in other jurisdictions, there is a section or Appendix included at the end which compares the outcome of the assessment with those of other Agencies (EPA, OK but perhaps others, as well).

For the same reasons, I think that the content of Sections 3.4.1.2.1.1, 3.4.2.2.1.2 and 3.4.1.2.2 (subsections 1, 2 and much of 3 and summary in 4) is not germane and rather detracts from the perception of impartiality of the TCEQ assessment (i.e., significant amounts of text dedicated to discrediting of the EPA analysis). Also, none of the arguments (summarized in the last para. of page 36 and overleaf to page 37) is particularly convincing in my view (e.g., margins between exposure in the occupational cohorts and environmental levels – these don't shed any light on the shape of the dose-response curve outside of the range of interpolation) nor do they add support to the defensibility of the TCEQ assessment. If objection to TCEQ not adopting the approach of U.S. EPA has been expressed in public comment, perhaps the lengthy text and analyses re same could be addressed in the responses there and not in the DSD.

Rather, analysis to support "ground truthing" or "reality checking" of the TCEQ approach is what's critical to defensibility of the assessment (only the text presented in the first and second paras. of page 49 and the associated content for the TCEQ estimates of the illustrative figures 8-12 on pages 52-55) and should stand on its own merits. I've made a number of suggestions re other "ground truthing" or comparisons which might contribute additionally to defensibility of the assessment (see response to Question 8).

Expert 5

As discussed in response to question 3, I believe that the case against the USEPA URF based on consideration of endogenous levels is not convincing. The DSD seems to make a lot out of the fact that the concentrations predicted by the USEPA model to correspond to $1e-6$ to $1e-4$ risk are in the range (even the lower part of the range) of the concentrations estimated to correspond to endogenous EtO levels. Given that lifetime lymphoid cancer probabilities are in the range of 0.02 to 0.03, as discussed above, the fact that adding exogenous exposures of roughly the same magnitude might bring the lifetime probabilities to 0.0201 to 0.0301 (for $1e-4$ added risk) does not seem incongruous to me. Obviously, such increases would be unmeasurable in practice. I do not understand why TCEQ finds such risks so hard to swallow.

Expert 6

The discussion of background endogenous levels and comparing them to endogenous levels produced by exogeneous levels is couched in terms of whether the endogenous levels produced by various exogeneous exposures are "biologically meaningful." This criterion does not necessarily take account of the very tiny risks ($10E-4$ to $10E-6$) that are considered de minimis. An exposure could be considered not biologically meaningful and still correspond to risks greater than the very tiny risks considered de minimis. The DSD, based on the Kirman and Hays (2017) estimates of the exogeneous exposures predicted to correspond to endogenous concentrations, calculates that the USEPA's URF of $7.1E-03$ per ppb suggests a background rate for lymphoid cancer caused by endogenous EtO equal to 46% of the

background rate of these tumors. But this is based on an (ADAF)-adjusted URF which itself is based on a statistical upper bound calculation of the URF, so the percentage could be quite a bit lower and not unreasonable. The endogenous levels of EtO in smokers appears to be very uncertain. Kirman and Hays (2017) present two estimates, one of which is only marginally above their estimate for non-smokers. The main point is that exposures that some consider “not biologically meaningful” may actually correspond to the tiny risks that are considered high enough to be of concern. The criterion upon which to judge these levels should be risk-based and not on whether they are considered “biologically meaningful.”

Charge Question 11

Please provide comments on the overall accuracy, objectivity, and transparency of the presentation of information in the revised DSD. Are the assumptions, data, and analyses described completely and clearly? Please identify any sections that need revision or improvement and describe in detail, to the extent possible, how they should be revised.

Expert 1

The summary of epidemiologic findings is presented clearly and objectively. I do not have sufficient expertise to comment on the risk assessment modeling.

Expert 2

This document will be widely read by a variety of audiences. I expect that it will be challenged in multiple venues, including the courts. For the DSD to have its intended impact, it is critical that its arguments and conclusions be presented in a clear and explicit manner. In its current state there is a lot of unclear, convoluted writing that detracts from message of the DSD.

I have noted on the text some edits to improve clarity, but I did not attempt much re-writing.

Note that many risk assessors and other scientists who are not mathematicians will be reading this document and attempting to understand the bases for conclusions. In the Appendices and elsewhere (see notes in the marked up DSD), I have indicated instances wherein a definition or plain language description would be useful.

It is often unclear as to which conclusions are those of TCEQ, US EPA, received wisdom, or some other source. It is important to understand which organization or person is making a statement, based on what rationale. Marginal comments or text edits are provided for some instances in the document (see notes in marked-up DSD).

The document is poorly organized, thus detracting from the strength of the arguments. Consider breaking the work into more, better organized chapters. We all acknowledge that hazard identification and dose response are integrated steps, but it may be easier on the reader if there is some better separation of descriptions.

I strongly suggest that the authors reconsider their use of Italics and other types of emphasis. These are more of a distraction to the reader than a help. It is not clear if there is a distinction between the meaning of Italics vs. bold type emphasis.

There is an unnecessary reliance on judgmental words (e.g. “remarkable”, “extraordinary”). These and other aspects of the writing contribute to a tone that is unnecessarily argumentative, if not defensive.

The overall effect is to raise suspicion in the mind of the reader as to the adequacy of arguments, which should stand on their own merits. In the mark up, I have noted some instances of this usage.

In the same vein, it is inappropriate and distracting to include arguments and observations regarding cancer incidence in a chapter on sources and uses of EtO (p. 12 par 1 and others). This has the overall effect of diluting the arguments that TCEQ is making on EtO carcinogenicity. This material may be better suited to a chapter on the rationale for TCEQ's review of EtO. Chapter 2 would benefit from reorganizing, as the material does not flow in a logical fashion.

The Chapter 3 section on WOE is very poorly organized, and it jumps among arguments. Moreover, some conclusions are declared in the absence of basis (e.g., p 18 last par: The sentence implies that animal data are not of any use in EtO cancer WOE, and this has not been supported).

Many arguments are not adequately described and supported (choice of MOA, descriptor of cancer classification). Other dead horses are continually flogged (estimates vs. observed cancers in NIOSH).

Even given an abbreviation / acronym list, the term should be written out at first use.

Note that "the agency" could refer to TCEQ, US EPA or another group. Please be explicit in the document as to who is saying what.

Expert 3

The document is quite thorough and fully explains the issues and basis for the judgments that were made. With regard to the persuasiveness of the presentation, there are two concerns, both editorial in nature:

1) The document is quite repetitive in revisiting the major concerns with the EPA approach to determining the URF. It would make it more digestible to see these laid out once in succinct form and then have sections of the report that methodically explain the basis for the overall comment. The many small sections and revisited topics make it somewhat difficult to read and to find where a particular point is fully developed.

2) There is a somewhat aggressive tone in attacking the EPA guidelines. This is related to the point above, repeatedly returning to the shortcomings of their model and its implications, somewhat defensive or even obsessive, suggesting outrage that goes beyond a careful scientific dissection which is present but pushed with unnecessary forcefulness. Throughout the report, there is a return to "pet peeves" and those interjections are both unnecessary and somewhat suggestive of a rigid point of view. A tone of calm examination and reasoning, including careful consideration of evidence that would lead to different judgments, would ultimately provide a more persuasive case for the decisions that were made.

Expert 4

As per the response to Question 10, it's exceedingly difficult to understand why so much effort and text in the assessment has been devoted to describing and countering an approach (US EPA, 2016) which appears to be inconsistent with TCEQ policy and methodological guidance (if I understand correctly), when the basis for discrepancy (i.e., outside the range of observation) is largely a matter of supposition. In my view, this detracts from the perceived impartiality and defensibility of the TCEQ assessment. This tone is rather prevalent, throughout, for which a few other examples are provided here.

My understanding based on the content of similar documentation in other jurisdictions is that outside of setting context (i.e., problem formulation), content is normally restricted to that relevant to the rationale for the approach and chemical specific evidence for the provision of health-based guidance on acceptable exposure levels.

For example, what is the relevance of inclusion of information in the paragraph at the top of page 7 related to undermining risk communication and societal consequences (Delete?). Similarly, there are some curious seemingly misplaced inclusions in Chapter 2, page 11, in a section entitled “Major Sources and Uses”, (second, third and fourth paragraphs) related to analyses of cancer incidence in Texas versus the U.S. and those of NATA ????.

As per comments in response to Question 10, suggested revised format to present a clearly delineated problem formulation to address the TCEQ mandate based on their methodological approach, technical experience and expertise and policy guidance is advised with comparison with other assessments relegated to a secondary role, perhaps in an Appendix and/or Responses to Public Comment.

Expert 5

As highlighted in response to other questions, I do have concerns about accuracy, objectivity and transparency, throughout the DSD. The text is overwrought in its repetitive, adversarial characterization of what others (primarily USEPA) have done. The text is made much harder to read with numerous (way TOO numerous) repetitions of the (often inaccurate) parenthetical descriptions or characterizations of a modeling approach or options. In particular, excessive repetition of the term “supra-linear” to characterize the USEPA model struck me as inflammatory.

That comment may in itself be seen as adversarial, and I apologize for that. But I just wanted to emphasize that I went into the review hoping, even expecting, to find a reasonable alternative presented and justified by a thorough and rigorous analysis. My expectations were not met.

Expert 6

My opinion is that DSD reflects a great deal of hard work and much thought. My biggest concern is with the characterization of the EPA model as supra-linear and that being the major reason given by the DSD for disregarding the EPA model. This error permeates the document and correcting it will entail a major revision. As noted in response to question 10 the DSD misinterprets the EPA conclusions about the low dose shape of the dose response in relation to endogenous EtO, and that needs to be corrected. I also think that, as noted in my response to question 10, the inferences drawn from the EtO levels in non-smokers and smokers have been overstated, given the uncertainty of these data, and that section should be revised. In the revision of this section the notion of whether certain exogenous exposures are “biologically meaningful” should be removed for the reason stated in my response to question 10. The focus of the evaluation of both the EPA model and the TCEQ model should be more on the low-dose fidelity, rather than the fit to the entire data set. In my response to the last question, I suggest some analyses that could be performed that might further illuminate the conclusions contained in the DSD.

Charge Question 12

The TCEQ solicited public comments on a June 2019 proposed DSD and has prepared a response to those comments (See Response to Public Comments Received on the Ethylene Oxide Draft Development Support Document, January 2020). Has the TCEQ appropriately addressed the critical scientific questions and issues raised by the public commenters in the Response to Comments and/or revised DSD? Are the responses to public comments presented clearly and completely? Please explain.

Expert 1

The TCEQ responses to the Public Comments regarding scientific and public comments appear to be appropriate and clearly stated. Further refinement of the DSD will undoubtedly add value to the responses.

Expert 2

The RTC document in most instances is clearly written. In some instances, the arguments are presented more clearly in the RTC than those in the DSD.

It appears to me that the major scientific questions and issues raised by the public commenters were addressed in the RTC, and the majority were addressed in the DSD, as well.

I suggest that TCEQ review the text in the DSD on the use of ADAF. I do not recall that the description of use and derivation of an adjusted URF completely addressed RTC p. 26, Comment 16.

I think additional clarifying text could be added to the DSD in response to RTC p.27 Comment 18. I don't recall that the document noted a reason for use of the 95% UCL on the calculated slope was to account for the uncertainty arising from lack of incidence data – or to make some sort of adjustment between cancer mortality and incidence.

I recommend that Figure 2 on page 50 of the RTC be reproduced in the DSD. This is in the context of my larger recommendation for expanded discussion of the genotoxicity data and the support (or lack thereof) for a mutagenic MOA.

I found RTC p. 63 point d. to be unclear.

Expert 3

The document provides a thorough and detailed response to the concerns that were raised.

Expert 4

As per responses to Questions 9 and 10, for the reasons mentioned there, I suspect that much of the text countering the U.S. EPA assessment should be included in the responses to public comment. This would enable appropriate focus within the DSD on the TCEQ rationale (the principal focus), which rather gets a bit lost, currently.

With respect to specific questions and issues raised by the public with the exceptions noted in response to the other questions, I believe that TCEQ has appropriately addressed the concerns in the Response to Comments and revised DSD.

Expert 5

The responses to comments are in the same vein as the DSD itself. I have already expressed my issues in relation to the DSD. The same applies generally to the responses to comments, though I acknowledge that it is often difficult to respond to comments succinctly and without repetition of the document itself.

Expert 6

In general, the TCEQ did a thorough and thoughtful job in responding to the many comments. However, many of TCEQ's comments relied on the faulty assumption that the EPA model is not acceptable because it is a supra-linear model. All of these comments are misleading and need to be corrected.

Charge Question 13

Please discuss any additional relevant comments or issues. Are there any additional questions or concerns that you would like fellow peer reviewers to address?

Expert 1

I have no additional comments or concerns at this time.

Expert 2

I have no additional relevant comments or issues at this time.

Expert 3

Perhaps the weakest part of the document pertains to the overreliance on statistical significance in judging the epidemiologic evidence. This is raised in regard to the validity of the EPA estimates on pages 34-37, throughout, and in table 5 and 6 (pp. 39-42). The claim is made that model assumptions are contradicted by the lack of statistically significant elevations in certain dose groups when it would have been expected based on the EPA model. It is reasonable to expect that a model derived directly from the data in an epidemiologic study such as the NIOSH study would generate estimates that are close to those observed within that study population. However, in evaluating the EPA model, the expectation for concordance with observations in the NIOSH cohort has to be tempered. While the overall shape of the dose-response curve should correspond, failure to observe statistically significant elevations in risk where the EPA model would predict elevated risks is inappropriate. Such an assessment does not address the overall shape of the dose-response curve to assess the compatibility of the EPA model and the NIOSH data. Furthermore, statistical significance is a function of effect size and precision, so that the model's predictions should be compared in a more global way with the observed data. "Non-significant" results may be entirely compatible with a prediction of elevated risk depending on the study size and precision of estimates. The judgment of concordance or lack of concordance between the model and the data needs to be more nuanced and less simplistic.

The use of ecologic data on pages 12-13 is an example of how the report creates the impression of a biased perspective, given that the authors recognize that the use of such studies to show EtO is not hazardous is fundamentally flawed. Such studies are far too crude to reveal much of anything about an environmental toxicant of this nature. It is noted that this is provided in response to public concerns, but if such studies had demonstrated elevated risks associated with the assigned exposures, it would be no more or less informative than it is with the absence of such associations. It would be preferable to note

that such questions have been raised and examined, but not to invoke those results as further evidence tempering the potential health effects of EtO exposure which they do not.

It is not clear from the report the extent to which the recommended limits are intended to err on the side of caution. Given the lack of clear guidance from empirical data, in making a series of judgments it does seem that a prudent approach is to take the uncertainty into account by building in some potential for error. It may be that the recommended URF does this but how the uncertainty was translated into prudence was not clear from the report.

Expert 4

The unpublished update of the UCC cohort (page 17, second para.) doesn't materially add to the content and since unverifiable at this stage, I'd suggest to delete.

Expert 5

I do not understand the distinctions being made in and around Figures 19-22 with respect to differences in estimated background obtained from the various models. I do not even understand what USEPA was referring to in the original notes to Figure 19. If the y-axis is RR, then I would think that the background assumed or estimated WOULD indeed be model-dependent. But the RR, being relative to whatever background is associated with the model in question, should be comparable across models (and data summaries), shouldn't it?

More generally, I would have been using the RR as a function of exposure as the "take-away" from the modeling, for each and every model. Those RR estimates could then be applied to a set of target-population background rates (e.g., life-table for the U.S., or maybe the Texas-specific, population). That way, one would not need to worry if the background rates are different, just that the RR as estimated is suitable for application to that target.

Expert 6

I have concerns about the TCEQ model and its fitting because it takes no account of the fact that the primary interest is in its predictions in the low dose range. One of the main points offered regarding its superiority over the EPA model is that overall (over all exposures) it provides a better description of the underlying data than the EPA model. It seems to me that this should not be a strong selling point since the focus should be on its fidelity in the low dose range, rather than its fit to the complete data set. In comparison, the EPA model, through its use of the spline, although it is fit to all the data, emphasizes fidelity to the low dose data. (However, in accomplishing this the spline dose response contains a very sharp bend at the spline point, so that the complete dose response is not very plausible.) These points leave me wondering whether a fit of some suitable dose-response model (perhaps a simple linear one) to a suitably selected low-dose subset of the data could provide a more appropriate answer. An analysis of this type is often used with animal data.

Both the EPA model and the TCEQ model are fit to the individual data, which is certainly preferred if all the data are completely trustworthy. However, in epidemiological studies exposure estimates, in particular, are often quite uncertain, and this is the case with the NIOSH data as well. So, I suggest it would be very worthwhile to check for outliers or highly influential data points, and to repeat analyses without these points. Similarly, I suggest that applying the models to "lightly categorized" data, using a fair number of categories, could be used as a possible check on results from the individual data.

It would be worthwhile to explain how the categorical data points were formed in Figures 20, 21 and 22 in the DSD. If that information is in the document, I missed it.

Finally, although the evidence for the carcinogenicity of EtO for breast cancer is not very strong, an estimate of its potential carcinogenicity could be included in TCEQ's analysis in the interest of not underestimating or appearing to underestimate the carcinogenicity of EtO. Alternatively, can an argument be made that inclusion of breast cancer would not change the regulatory impact of TCEQ's analysis? Based on EPA's analysis it appears that this might be the case.

References Cited

- Coggon, H, Harris, E, Poole, J, Palmer, K. 2004. Mortality of workers exposed to ethylene oxide extended follow up of a British cohort. *Occup Environ Med.* 61:358-362.
- Crump KS. 1984. A New Method for Determining Allowable Daily Intakes. *Fund Appl Toxicol.* 4:854-871.
- Crump KS. 1995. Calculation of Benchmark Doses from Continuous Data. *Risk Analysis* 15(1):79-89.
- Crump KS. 2011. Use of threshold and mode of action in risk assessment. *Critical Reviews in Toxicology*, 41(8), 637–650. <https://doi.org/10.3109/10408444.2011.566258>.
- Crump KS. 2017. Bogen’s Critique of Linear-No-Threshold Default Assumptions: Perspective. *Risk Analysis*, 37(10):1802-1807. doi:10.1111/risa.12748.
- Crump, KS, & Allen, BC. 1985. Methods for Quantitative Risk Assessment Using Occupational Studies. *The American Statistician* 39(4):442-450.
- Crump KS, Hoel DG, Langley CH, Peto R. 1976. Fundamental carcinogenic processes and their implications for low dose risk assessment. *Cancer Research*, 36(9 pt.1):2973–2979.
- Crump KS, Bussard DA, Chen C, Jinot J, Subramaniam R. 2014. The “bottom-up” approach does not necessarily bound low-dose risk. *Regul Toxicol Pharmacol*, 70(3): 735–736. <https://doi.org/10.1016/j.yrtph.2014.10.008>.
- Davis, J. Gift, J. Zhao, Q. 2010. Introduction to benchmark dose methods and U.S. EPA's Benchmark Dose Software (BMDS) version 2.1.1. *Toxicol Appl Pharm.* 254. 181-91. 10.1016/j.taap.2010.10.016
- Donner, EM, Wong, BA, James, RA, Preston, RJ. 2010. Reciprocal translocations in somatic and germ cells of mice chronically exposed by inhalation to ethylene oxide: Implications for risk assessment. *Mutagenesis* 25:49-55.
- Health Canada. 2001. Priority Substance List Assessment Report: Ethylene Oxide. Available at: https://www.canada.ca/content/dam/hc-sc/migration/hc-sc/ewh-semt/alt_formats/hecs-sesc/pdf/pubs/contaminants/psl2-lsp2/ethylene_oxide/oxyde_ethylene_oxide-eng.pdf
- International Agency for Research on Cancer (IARC). 2008. IARC monographs on the evaluation of carcinogenic risks to humans: 1,3-butadiene, ethylene oxide and vinyl halides (vinyl fluoride, vinyl chloride and vinyl bromide). Lyon, France.
- International Agency for Research on Cancer (IARC). 2012. IARC monographs on the evaluation of carcinogenic risks to humans, No. 100F. Lyon, France.
- International Agency for Research on Cancer (IARC). 2019. Tumour site concordance and mechanisms of carcinogenesis. Part 3, Chapter 21: Analysis of tumour site concordance. IARC Scientific Publication No. 165. Lyon, France.
- Kirkeleit, J, Riise, T, Bjørge, T, Christiani, DC. 2013. The healthy worker effect in cancer incidence studies. *Am J Epidemiol.* 2013; 177(11): 1218-1224.

- Kirman, C, Hays, S. 2017. Derivation of endogenous equivalent values to support risk assessment and risk management decisions for an endogenous carcinogen. *Regulatory Toxicology and Pharmacology*. 91:165-172.
- Marsh, G.M., Keeton, K.A., Riordan, A.S., Best, E.A., Benson, S.M. 2019. Ethylene oxide and risk of lympho-hematopoietic cancer and breast cancer: a systematic literature review and meta-analysis. *Int. Arch. Occup. Environ. Health* 92, 919–939.
- Meek, M.E. 2017. AOPs in hazard characterization for human health. *Current Opinion in Toxicology* 3:80–86.
- Meek, M.E., Boobis, A., Cote, I., Dellarco, V., Fotakis, G., Munn, S., Seed, J. and Vickers, C. 2014a. New developments in the evolution and application of the WHO/IPCS framework on mode of action/species concordance analysis. *J. Appl. Toxicol.*, 34: 1-18. doi:10.1002/jat.2949
- Meek, M.E., Palermo, C.M., Bachman, A.N., North, C.M. and Lewis, R.J. 2014b. Mode of action human relevance (species concordance) framework: Evolution of the Bradford Hill considerations and comparative analysis of weight of evidence. *J. Appl. Toxicol.*, 34: 595-606. doi:10.1002/jat.2984
- Mikoczy, Z, Tinnerberg, H, Björk, J, Albin, M. 2011. Cancer incidence and mortality in Swedish sterilant workers exposed to ethylene oxide updated cohort study findings 1972-2006. *Int. J. Environ. Res. Public Health*. 8:2009-2019.
- Norman, S, Berlin, J, Soper, K, Middendorf, B, Stolley, P. 1995 Cancer incidence in a group of workers potentially exposed. *International Journal of Epidemiology*. 24(2):276-284.
- Organisation for Economic Development (OECD). 2018. Users' Handbook supplement to the Guidance Document for developing and assessing Adverse Outcome Pathways, *OECD Series on Adverse Outcome Pathways*, No. 1, OECD Publishing, Paris, <https://doi.org/10.1787/5jlv1m9d1g32-en>.
- Steenland, K., Whelan, E., Deddens, J., Stayner, L., Ward, E. 2003. Ethylene oxide and breast cancer incidence in a cohort study of 7576 women (United States). *Cancer Causes Control* 14, 531–539.
- Steenland, K., Stayner, L., Deddens, J. 2004. Mortality analyses in a cohort of 18 235 ethylene oxide exposed workers: follow up extended from 1987 to 1998. *Occup. Environ. Med.* 61, 2–7.
- Swenberg, J.A., Fryar-Tita, E., Jeong, Y.C., Boysen, G., Starr, T., Walker, V.E., Albertini, R.J. 2008. Biomarkers in toxicology and risk assessment: Informing critical dose-response relationships. *Chem. Res. Toxicol.* 21, 253–265.
- TCEQ. 2015. Guidelines to develop toxicity factors. RG-442: Texas Commission on Environmental Quality (TCEQ).
- United States Environmental Protection Agency (US EPA). 2005. Guidelines for Carcinogen Risk Assessment. (EPA/630/P-03/001B). Washington, D.C.
- United States Environmental Protection Agency (US EPA). 2016. Evaluation of the inhalation carcinogenicity of ethylene oxide (CASRN 75-21-8): In support of summary information on the Integrated Risk Information System (IRIS). EPA/635/R-16/350Fa. Washington, DC, U.S. Environmental Protection Agency, Office of Research and Development, National Center for Environmental Assessment.

Vincent, MJ, Kozal, JS, Thompson, WJ, et al. 2019. Ethylene oxide: cancer evidence integration and dose-response implications. *Dose-Response: An International Journal* October-December 1-17 (<https://journals.sagepub.com/doi/10.1177/1559325819888317>).

Wong, O, Trent, L. 1993. An epidemiological study of workers potentially exposed to ethylene oxide. *British Journal of Industrial Medicine*. 50:308-316.

Appendix A – Peer Review Organization

The RSC organized this peer review and is responsible for managing all aspects of the process, including selection of the reviewers, evaluation of potential conflicts of interest, development of the charge questions, distribution of the assessment document, collection and review of each expert's written comments, and collation of all comments into this report. RSC uses best practices for peer review that reflect guidance from the Office of Management and Budget, the National Academy of Sciences, the U.S. Environmental Protection Agency (US EPA) and others.

Selection of Reviewers. RSC reviewed the draft document and discussed with TCEQ the types of scientific and technical expertise needed for this review. Specifically, the following types of expertise were identified: epidemiology, carcinogenicity, toxicology, dose-response modelling, mode of action, and weight of evidence. RSC identified candidate reviewers with the appropriate expertise using a number of strategies and tools, including an internal database of experts, and literature and Internet searches for those publishing on EtO and relevant issues. We sought candidates with diverse backgrounds and perspectives (e.g., government, academia, industry, consulting) to provide a comprehensive and technically-sound review.

Prior to contacting potential candidates, we considered what would constitute a conflict of interest or a bias that might prevent an expert from providing an objective evaluation of the TCEQ document. RSC identified situations and conditions that may be considered potential conflicts of interest (COI) for the peer reviewers and developed a COI questionnaire to screen the candidates. The following is a list of parties that we expect would have a financial interest in, or have a pre-existing position on, the TCEQ DSD or the outcome of this review:

- Texas Commission on Environmental Quality (TCEQ)
- Companies that manufacture, use, or import ethylene oxide
- Medical sterilization companies and facilities
- Relevant industry trade associations
- U.S. Environmental Protection Agency
- Those who provided public comments on TCEQ's proposed DSD

The initial candidate list was screened to remove those who met any of the following criteria:

- Individuals who work for TCEQ
- Individuals who authored or contributed to the TCEQ DSD
- Individuals who authored or contributed to EPA's IRIS assessment for EtO
- Individuals who are employed by companies that manufacture, import, or use EtO
- Individuals who published on EtO, if the funding for the work came from any of the above interested parties

Selected candidates were contacted to determine their qualifications and whether they were interested in being a peer reviewer. RSC evaluated the candidates' technical/scientific background and their relevant knowledge and experience in one or more key subject areas for the assessment.

Qualified and interested candidates were asked about funding and/or affiliations that they may have with interested parties and current or previous work on EtO. They were asked to complete a COI questionnaire that included a series of questions exploring the potential for financial conflicts of interest and circumstances that might result in a lack of impartiality, which might prevent an expert from

providing an independent scientific opinion on the EtO assessment. Candidates were asked to consider their financial and other relationships with the list of interested parties and to report on relationships they may have with these parties or any others that may have an interest in the EtO risk assessment.

The RSC selected a group of six experts that provide a diversity and balance of relevant expertise and backgrounds. The group included at least two individuals who are expert in each of these key areas: environmental epidemiology; cancer dose response modeling; and, cancer risk assessment/toxicology/mode of action.

The list of experts was shared with TCEQ to allow the Commission to identify any reviewers that TCEQ thought may have a potential conflict of interest or may be unqualified. TCEQ raised a question about one expert, noting a published opinion regarding some US EPA policies. RSC discussed this with the expert and concluded that the individual's opinions as expressed in the blog would not interfere with that individual's ability to provide an objective review of the TCEQ EtO assessment.

The selected expert peer reviewers for this review are listed below. Their affiliations are provided for identification purposes only. Biographical sketches and relevant information with regard to conflict of interest are found in Appendix C.

- Bruce Allen, M.S., Biomathematician, independent consultant
- Kenny Crump, Ph.D., Biomathematician, private consultant, retired
- Harvey Checkoway, Ph.D., Professor of Epidemiology, Department of Family Medicine and Public Health and Department of Neurosciences, University of California, San Diego
- Bette Meek, Ph.D., Associate Director of Chemical Risk Assessment, McLaughlin Centre for Risk Science, University of Ottawa
- David Savitz, Ph.D., Professor of Epidemiology and Interim Chair, Department of Epidemiology in the School of Public Health, and Obstetrics and Gynecology and Pediatrics in the Alpert Medical School, Brown University
- Rita Schoeny, Ph.D., Consultant in risk assessment and science policy, retired from US EPA

Collectively, the panel of experts has backgrounds in consulting, academia, research, and government. Because EtO is so widely used in the the chemical and sterilization industry, and industry trade groups have expressed opinions in public comments on the TCEQ proposed DSD, we were not able to find experts working for industry that were not employed by interested parties. Likewise, major environmental non-profits and health advocacy groups have voiced their opinions on the DSD in public comments to TCEQ, which precluded selection of experts affiliated with those groups.

The peer reviewers were asked at each teleconference whether they had any changes or updates to their conflict of interest information or questions about other reviewers' information. No experts had changes or questions.

Development of Charge. TCEQ identified a number of questions and issues on which it sought expert opinions. RSC reviewed the TCEQ suggestions and the DSD document and public comments, and drafted a list of charge questions. The draft questions were sent to TCEQ for comment and input. RSC considered TCEQ's input on the charge questions, and independently decided on the final wording of

the questions. Both focused and open-ended questions were included to provide the experts the opportunity to identify and raise any issues that they felt were important to be addressed.

A copy of the charge and instructions for peer reviewers is found in Appendix B.

Review Package. The review package was emailed to experts on February 2, 2020. It included the following:

- TCEQ EtO assessment (*Revised Draft EtO DSD 1-31-20*)
- TCEQ response to public comments document (*RTC for Revised Draft EtO DSD 1-31-10*).
- EtO Panel biographical sketches and conflict of interest information
- Charge Questions
- Key References (list included in Charge question document)

Experts were given five weeks to review the draft documents and prepare their initial responses to the charge questions. Experts were asked to contact the RSC for any help they may need obtaining additional references or information.

Teleconference. A teleconference was held on February 11, 2020 with TCEQ staff, RSC staff, and the experts. The purpose of the teleconference was to provide the experts with background and context for the TCEQ DSD, to review the charge questions to make sure they are clear and complete, and to answer questions from the experts regarding the peer review process or materials. After introductions, the experts were asked if they had any questions, comments or revisions regarding conflict of interest; none did. TCEQ staff briefly reviewed the history of the EtO DSD development and explained that the DSD will be used for facility permits and to evaluate air monitoring data. They emphasized that they are not asking experts to choose between EPA (2016) and the TCEQ DSD, but are seeking comments regarding the scientific defensibility of the TCEQ assessment. TCEQ explained that all the public comments received were considered in the revisions made to the DSD document the experts are reviewing.

In response to experts' questions, RSC staff clarified that the names of the experts will be listed in the final report, along with biographical sketches and conflict of interest information, but that the individual comments will be anonymous. They also noted that after the peer review, TCEQ will prepare a response to peer review comments document, which will be publicly available.

Experts raised the following questions regarding the Charge Questions:

- Question 5 – An expert asked TCEQ to identify the part of the DSD to which the question of mortality vs. incidence is directed. TCEQ noted that this was an issue raised in public comments and is more a general question of which approach is appropriate.
- Question 7 – An expert asked if TCEQ is looking for comments on the supralinear model specific to EtO, or more general comments. RSC clarified that EtO-specific comments are sought.
- Question 13 – An expert suggested that any new questions that experts may have for one another should be raised early in the process so there is time to assign the question to a specific expert for response.

The experts had no suggestions for additions or revisions to the charge questions.

TCEQ noted that they have made a small revision to the DSD and a small revision to the response to comments document and the RSC sent the revised documents to the experts.

Reviewers' Initial Comments and Draft Report. The experts sent their initial responses to the charge questions to RSC. The RSC staff reviewed the experts' responses to ensure they were complete and clear. We contacted experts as needed to clarify a few minor items (e.g., references).

RSC collated the initial expert comments into a draft report, which was organized by charge question so that one can see all reviewers' comments and opinions on each charge question together. The experts were randomly assigned a number to identify their responses, allowing the specific comments to remain anonymous. This was done to make the experts feel more comfortable in expressing their candid opinions. The draft report also included a copy of the charge questions and instructions to experts, and expert biographical sketches and conflict of interest information.

The draft report was sent to TCEQ and to the experts for their review and consideration prior to a teleconference on April 8, 2020. The purpose of the teleconference was to provide the opportunity for reviewers and TCEQ to seek any needed clarification regarding the experts' responses to the charge questions; however consensus opinions were not sought. After the teleconference, the experts were given the opportunity to revise their written responses to the charge questions and their final responses were collated here in this final report of the peer review.

This final report constitutes the record of this peer review. A copy of the final report will be made available to the public on the RSC web site (<https://med.uc.edu/eh/centers/rsc/peer-review>).

Appendix B - Charge Questions and Instructions

Charge Questions

TCEQ Ethylene Oxide Development Support Document (DSD)

Background and Timeline

The following two paragraphs were provided by TCEQ to provide background for the experts:

Ethylene oxide (EtO) is used as a chemical intermediate in the manufacture of ethylene glycol (antifreeze), polyester, detergents, polyurethane foam, solvents, medicine, adhesives, and other products. Relatively small amounts of EtO are used in sterilization of surgical equipment and plastic, as a fumigant, and as a sterilant for food (spices) and cosmetics. In 2018, EtO was being produced in the US at 15 facilities in 11 locations by nine companies. In the US, EtO is primarily produced in Texas and Louisiana. IARC has designated EtO as a group 1 human carcinogen (IARC, 2012). Between October 1, 2018 and March 31, 2019, the EPA conducted air monitoring for EtO in various locations in the United States, and found that the levels of EtO concentrations that are considered to be “urban background” are in the range of 0.1 – 0.2 ppb (<https://www.epa.gov/hazardous-air-pollutants-ethylene-oxide/ethylene-oxide-data-summary-national-air-toxics-trends>). In regard to longer-term levels around EtO-emitting facilities, as an example, the mean and 95th percentile modeled 5-year concentrations for one sterilizer facility were ≈0.17 and 1.6 ppb, respectively ([https://www.atsdr.cdc.gov/HAC/pha/sterigenic/Sterigenics International Inc-508.pdf](https://www.atsdr.cdc.gov/HAC/pha/sterigenic/Sterigenics_International_Inc-508.pdf)).

In early 2017, as part of a standard yearly review of newly-derived toxicity factors, the TCEQ Toxicology Division reviewed the EPA’s cancer-based toxicity factor derivation for EtO (finalized in 2016) to determine if we would provisionally adopt the EPA’s number for use in deriving protective concentration levels (PCLs) for the Texas Risk Reduction Program (TRRP). In March 2017 the division decided that, instead of adopting the EPA’s EtO toxicity factor, we would derive an interim EtO toxicity factor for the agency’s use in the remediation program with a plan to conduct a complete evaluation of EtO inhalation carcinogenicity for use in both air permitting and remediation. The TCEQ decided to complete this thorough evaluation because EtO is emitted in Texas and has been determined to be a carcinogen. In August of 2017, the TCEQ announced a 90-day public information request for scientific information about EtO that may be of use in the TCEQ’s review. The TCEQ then completed a systematic review and dose-response assessment of EtO carcinogenicity and released the draft Development Support Document (DSD) on June 28, 2019 for public comment. The public comment period ended in late September 2019 and the TCEQ undertook a review and response to the public comments and scientifically justified revisions to the draft DSD. The revised DSD and response to comments are the documents that are under review in this peer review.

The purpose of this peer review is to provide TCEQ with expert review of the development of their inhalation cancer unit risk factor (URF) for EtO to determine if the work is scientifically adequate and

appropriate for estimating cancer risk at ambient (low-level) concentrations. Experts are also asked to comment on the clarity and transparency of the documentation.

Review Package

The following documents and materials have been provided to the peer reviewers:

- Revised Ethylene Oxide DSD (Revised Draft EtO DSD 1-31-20)
- TCEQ Response to Public Comments (RTC for Revised Draft EtO DSD 1-31-20)
- Copies of all substantive public comments received on the proposed DSD (14 documents)
- Key references
 - IARC. 2019. Tumour site concordance and mechanisms of carcinogenesis. Part 3, Chapter 21: Analysis of tumour site concordance. IARC Scientific Publication No. 165.
 - Kirkeleit, J., Riise, T., Bjørge, T., Christiani, D.C., 2013. The healthy worker effect in cancer incidence studies. *Am. J. Epidemiol.* 177, 1218–1224.
 - Kirman, C.R., Hays, S.M., 2017. Derivation of endogenous equivalent values to support risk assessment and risk management decisions for an endogenous carcinogen: Ethylene oxide. *Regul. Toxicol. Pharmacol.* 91, 165–172.
 - Marsh, G.M., Keeton, K.A., Riordan, A.S., Best, E.A., Benson, S.M., 2019. Ethylene oxide and risk of lympho-hematopoietic cancer and breast cancer: a systematic literature review and meta-analysis. *Int. Arch. Occup. Environ. Health* 92, 919–939.
 - SAB (Science Advisory Board). 2007. Review of Office of Research and Development (ORD) draft assessment entitled "Evaluation of the carcinogenicity of ethylene oxide". Washington, DC: U.S. Environmental Protection Agency, Science Advisory Board. [http://yosemite.epa.gov/sab/sabproduct.nsf/368203f97a15308a852574ba005bbd01/5D661BC118B527A3852573B80068C97B/\\$File/EPA-SAB-08-004-unsigned.pdf](http://yosemite.epa.gov/sab/sabproduct.nsf/368203f97a15308a852574ba005bbd01/5D661BC118B527A3852573B80068C97B/$File/EPA-SAB-08-004-unsigned.pdf)
 - SAB (Science Advisory Board). 2015. Science Advisory Board Review of the EPAs evaluation of the inhalation carcinogenicity of ethylene oxide: Revised external review draft - August 2014 [EPA Report]. (EPA-SAB-15-012). Washington, DC: U.S. Environmental Protection Agency, Science Advisory Board.
 - [https://yosemite.epa.gov/sab/sabproduct.nsf/fedrgstr_activites/BD2B2DB4F84146A585257E9A0070E655/\\$File/EPA-SAB-15-012+unsigned.pdf](https://yosemite.epa.gov/sab/sabproduct.nsf/fedrgstr_activites/BD2B2DB4F84146A585257E9A0070E655/$File/EPA-SAB-15-012+unsigned.pdf)
 - Steenland, K., Whelan, E., Deddens, J., Stayner, L., Ward, E., 2003. Ethylene oxide and breast cancer incidence in a cohort study of 7576 women (United States). *Cancer Causes Control* 14, 531–539.
 - Steenland, K., Stayner, L., Deddens, J., 2004. Mortality analyses in a cohort of 18 235 ethylene oxide exposed workers: follow up extended from 1987 to 1998. *Occup. Environ. Med.* 61, 2–7.
 - Swaen, G.M.H., Burns, C., Teta, J.M., Bodner, K., Keenan, D., Bodnar, C.M., 2009. Mortality study update of ethylene oxide workers in chemical manufacturing: A 15 year update. *J. Occup. Environ. Med.* 51, 714–723.
 - Swenberg, J.A., Fryar-Tita, E., Jeong, Y.C., Boysen, G., Starr, T., Walker, V.E., Albertini, R.J., 2008. Biomarkers in toxicology and risk assessment: Informing critical dose-response relationships. *Chem. Res. Toxicol.* 21, 253–265.
 - TCEQ, 2015. RG-442 Guidelines to Develop Toxicity Factors.

- US Environmental Protection Agency. 2016. Evaluation of the Inhalation Carcinogenicity of Ethylene Oxide, in Support of Summary Information on the Integrated Risk Information System. EPA/635/R-16/350Fa.
https://cfpub.epa.gov/ncea/iris/iris_documents/documents/toxreviews/1025tr.pdf
- US Environmental Protection Agency, ethylene oxide derivation, previous drafts – choose Chemical Documents:
https://cfpub.epa.gov/ncea/iris2/chemicalLanding.cfm?substance_nmbr=1025#tab-6
- Valdez-Flores, C., Sielken, R.L., Teta, M.J., 2010. Quantitative cancer risk assessment based on NIOSH and UCC epidemiological data for workers exposed to ethylene oxide. Regul. Toxicol. Pharmacol. 56, 312–20. And supplemental material.
- Valdez-Flores, C., Sielken, R.L., 2013. Misinterpretation of categorical rate ratios and inappropriate exposure-response model fitting can lead to biased estimates of risk: Ethylene oxide case study. Regul. Toxicol. Pharmacol. 67, 206–214.
- Vincent, M.J., Kozal, J.S., Thompson, W.J., Maier, A., Dotson, G.S., Best, E.A., Mundt, K.A., 2019. Ethylene Oxide: Cancer Evidence Integration and Dose–Response Implications. Dose-Response 17, 1–17. And supplemental material.

Copies of additional references or supporting documentation will be provided to reviewers upon request (email Brian Palmer at palmerbr@ucmail.uc.edu.)

Instructions

For each response (including the Yes/No questions), please explain your reasoning and considerations, discuss scientific support for your comments and opinions, and identify the sources you consulted to construct your response. Please address each charge question by adding your answers to this Word document; indicate page, paragraph, and line number, where applicable. In addition, you may suggest editorial revisions to the DSD document directly by using track changes. If you feel a question is outside your area of expertise, please indicate “Outside my area of expertise.”

Due Date - Your draft written review should be returned to Jacqueline.patterson@uc.edu no later than March 11, 2020. You will have the opportunity to finalize your comments after the second conference call, which will be held around April 8, 2020

Charge Questions

1. The TCEQ conducted a systematic review of the literature relevant to the derivation of an inhalation unit risk factor for ethylene oxide (EtO) (see DSD Appendix 1). Are you aware of any additional literature or studies that should be considered and if so, how might they impact the assessment?
2. The TCEQ adopts the EPA conclusion that the weight of the evidence supports a direct-acting mutagenic mode of action (MOA) for EtO carcinogenicity (DSD Sections 3.3 and 3.3.1). Section 3.3.1 of the DSD presents summary information from the EPA (EPA Section 3.4.3) relevant to the MOA determination. Do you agree with the MOA determination? Please explain.

3. The TCEQ adopts EPA's MOA analysis (DSD Section 3.3.1) and considers MOA as information relevant to the likely or expected shape of the dose-response (DSD Sections 3.4.1 and 3.4.1.1) as specified by the TCEQ guidelines for developing toxicity factors (TCEQ, 2015). What is your opinion on whether and how the MOA should inform the likely or expected shape of the dose-response curve, overall and in the low-dose range (e.g., at environmentally-relevant concentrations); and whether and how the MOA should inform the choice of dose-response model for estimating human carcinogenicity risk? Please comment on TCEQ's reasoning on the implications of the MOA for the shape of the dose-response and its relative importance amongst their other model choice considerations (summarized in DSD Section 3.4.1.4.2). Are the TCEQ conclusions concerning implications of the MOA scientifically defensible?
4. The TCEQ conducted an evaluation of EtO's carcinogenic classification (DSD Section 3.3.2), and also evaluated breast cancer risk in humans as a potential cancer endpoint (DSD Appendix 6; Response to Dr. Kyle Steenland, Comment 1 in Response to Public Comments Document). What is your characterization of the overall weight of the evidence for or against EtO increasing the risk of breast cancer in humans at occupational concentrations (past or present) and at environmentally-relevant concentrations?
5. While it is in the interest of public health to protect against cancer *incidence*, available epidemiological studies often only provide cancer *mortality* data for dose-response modeling. What is your opinion on the accuracy of using a dose-response model based on cancer mortality data (e.g., lymphoid cancer mortality) to predict cancer incidence (e.g., lymphoid cancer incidence)?
6. The TCEQ's DSD discusses a problem with key USEPA AIC and p-value calculations used as criteria in determining model fit, and the TCEQ recalculated these values (DSD Section 3.4.1.3 and Appendix 4). Please explain what you think the appropriate approach should be for accounting for the number of estimated parameters in the modeling and the associated calculation of the AIC and p-values. Given that appropriate AIC and p-values are available for models fit to *individual* data, what role should visual fit to *categorical* estimates play in model selection (Response to University of California at San Francisco, Comment 6 in Response to Public Comments Document)?
7. Please comment on the biological and mechanistic support for and against use of an overall supralinear model to estimate risk of lymphoid cancer from exposure to EtO at occupational levels and at environmentally-relevant concentrations.

8. As summarized in DSD Section 3.4.1.4.2, the TCEQ used MOA, model predictiveness reality checks (both for the NIOSH cohort and the general population), biological plausibility, and statistical model fit criteria for model selection. Have these considerations been clearly described and are they scientifically appropriate given the available data?
9. In DSD Sections 3.4.1.4, 3.4.1.5, and 3.4.1.6, the TCEQ describes their modeling choices and assumptions, and calculates an inhalation unit risk factor (URF), ultimately applying age-dependent adjustment factors (ADAFs) in DSD Section 3.4.2. Do you disagree with any of the modeling choices and assumptions or calculations made by TCEQ in the dose-response assessment? Please discuss any issues or concerns you have with the inhalation URF derivation.
10. Based on biomarker data, various sections of the DSD (e.g., Section 3.4.1.2.1, Section 3.4.1.4.2 number "4.", second to the last paragraph of Section 3.4.1.6.2) discuss air concentrations corresponding to endogenous and background EtO levels and also compare these levels to acceptable air concentrations derived from URFs (either the TCEQ's or EPA's). Such a discussion is also included in the Response to Public Comments document (e.g. Response to Dr. Kyle Steenland, Comment 3). Please comment on whether the information and context provided by the discussion of endogenous/background EtO levels is clear and is scientifically appropriate.
11. Please provide comments on the overall accuracy, objectivity, and transparency of the presentation of information in the revised DSD. Are the assumptions, data, and analyses described completely and clearly? Please identify any sections that need revision or improvement and describe in detail, to the extent possible, how they should be revised.
12. The TCEQ solicited public comments on a June 2019 proposed DSD and has prepared a response to those comments (See Response to Public Comments Received on the Ethylene Oxide Draft Development Support Document, January 2020). Has the TCEQ appropriately addressed the critical scientific questions and issues raised by the public commenters in the Response to Comments and/or revised DSD? Are the responses to public comments presented clearly and completely? Please explain.
13. Please discuss any additional relevant comments or issues. Are there any additional questions or concerns that you would like fellow peer reviewers to address?

Appendix C –Biosketches and Conflict of Interest Screening

To facilitate the identification and evaluation of potential conflict of interest and bias situations for the peer review candidates, the University of Cincinnati Risk Science Center (RSC) identified a list of potentially affected or interested parties and sectors for this peer review of an ethylene oxide (EtO) cancer dose-response assessment development support document (DSD).

The candidates were asked to consider their financial and other relationships with a list of interested parties when completing a conflict of interest questionnaire and to report any relationships they may have with these parties. The list included the Texas Commission on Environmental Quality (TCEQ); the authors and contributors to the TCEQ document; the U.S. Environmental Protection Agency (US EPA); companies that manufacture, import, or use EtO; sterilization companies; relevant trade associations; and those parties that submitted public comments on the proposed TCEQ DSD. The candidates were also questioned about current and past activities on ethylene oxide.

The RSC has evaluated the information utilizing guidance and procedures developed by the U.S. National Academy of Sciences and the US EPA for committee membership and external peer review, respectively. The RSC has determined that none of the experts has a conflict of interest with this peer review, nor do any of the experts have anything in their background or experience that would interfere with their abilities to provide an impartial, technically sound, and objective review of the subject matter.

In the interest of transparency, relevant information is reported below for each candidate. In addition, relevant information regarding conflict of interest for the University of Cincinnati RSC (the group that organized and conducted the peer review under contract to TCEQ) is also provided below.

Ethylene Oxide Peer Review Panel Members

RSC independently selected the following six experts to provide independent peer review of the TCEQ EtO DSD. Each has been screened for conflict of interest. None of the selected experts has a conflict of interest with the review of this document.

Bruce Allen

Bruce C. Allen is a biomathematician and works as an independent consultant. Previously Mr. Allen held manager and senior scientist positions with Environ International; ICF Consulting, KS Crump Group; and RAS Associates. He received his M.S. in Biomathematics with a Statistics minor from North Carolina State University. Mr. Allen's areas of expertise are dose response analysis, biologically-motivated modeling, and statistics, particularly applied to human health risk assessment. His primary interest is in the quantitative aspects of risk analysis, including dose-response analysis; statistical appraisal of data, models, and modeling results; and with developing rigorous approaches to decision making in risk assessment contexts. In particular, Mr. Allen has conducted research to study dose-response modeling approaches for developmental toxicants and cancer dose-response, and the estimation of risks from epidemiological data. He has over 50 peer-reviewed publications on quantitative risk assessment and is a frequent peer reviewer of risk assessment documents. Over the past five years, Mr. Allen has consulted on a paid basis for the U.S. EPA (through contractors Lockheed Martin, CSRA, and GDIT), the University of Cincinnati RSC, Toxicology Excellence for Risk Assessment, Ramboll Corporation, ICF Consulting, and 3M.

Mr. Allen recently was a co-author of a paper³ published in 2017 on dose and temporal evaluation of ethylene oxide-induced mutagenicity in lungs of mice. The paper was based on research performed under a CRADA with the FDA, with funding from Ethylene Oxide and Derivatives Producers' Association and Lower Olefins Sector Group of the European Chemical Industry Association (CEFIC). This government/industry collaborative project addressed the MOA for EtO-induced lung tumors in mice, which is not an issue relevant to the TCEQ ethylene oxide assessment. Mr. Allen (as a sub-contractor to TERA) contributed dose-response analysis to the project and publication. Mr. Allen's work on this project does not constitute a conflict of interest and will not interfere with Mr. Allen's ability to be impartial with this peer review, because the work is completed, no other work is anticipated, and it is not relevant to the current assessment. For the U.S. EPA, Mr. Allen is currently working on several projects (as a subcontractor to a government contractor), including dose-response analyses for arsenic and development of benchmark dose software. Since 2015 he has worked for 3M (a company that produces sterilization equipment using EtO) on dose-response and pharmacokinetic modeling of PFAS and related compounds (some of that work done as a subcontractor through a consulting firm). These projects are unrelated to EtO, do not create a conflict interest, and will not interfere with Mr. Allen's ability to provide an impartial, technically sound, and objective review of the subject matter.

Harvey Checkoway

Harvey Checkoway has been a Professor since 2013 and Vice Chair for Research since 2014 at the University of California, San Diego (UCSD), in the Department of Family Medicine and Public Health. He has had a joint appointment as a Professor in the UCSD Department of Neurosciences since 2015. Previously he was a professor in the Departments of Environmental and Occupational Health Sciences and Epidemiology at the University of Washington (1987-2013) and an assistant professor (1979-86), and associate professor (1987) at the University of North Carolina. He earned his Ph.D. in epidemiology from the University of North Carolina and an M.P.H. from Yale University. His research has focused on epidemiologic investigations of occupational and environmental risk factors for chronic diseases, especially cancers and neurodegenerative disorders. His research has included assessments of biomarkers in occupational epidemiological investigations and methodological research on study design and research validity considerations, as well as evaluation of the quality of epidemiological research data for applications to risk assessments. Dr. Checkoway has served on many advisory committees and peer review panels for state, federal, and international government agencies, unions, and private companies. For the National Academy of Sciences he has served on committees addressing military personnel at atmospheric tests of nuclear weapons, low-dose ionizing radiation, IRIS formaldehyde risk assessment, and asbestos. Dr. Checkoway also served as a Visiting Scientist at the International Agency for Research on Cancer (IARC) in 2006, and for 1-3-month periods as a Visiting Professor in the Department of Occupational and Environmental Medicine at the University of Turin in 2017, and as a Visiting Scientist at Bologna University Institute of Advanced Studies in 2019. Dr. Checkoway is the lead author of a major textbook on research methods in occupational epidemiology and is a journal editor and reviewer. He has over 250 peer-reviewed publications, and has written 25 book chapters. During the last five years Dr. Checkoway has received grants or funding from NIEHS, NIOSH, National Cancer Institute, National Multiple Sclerosis Society, National Institute for Neurological Diseases and Stroke, National Health Lung and Blood Institute, Alpha Foundation for the Improvement of Mine Safety and Health, and the National Institute of Aging. As a consultant, he has also received funding from the

³ Manjanatha, Mugimane G; Shelton, Sharon D; Chen, Ying; Parsons, Barbara L; Myers, Meagan B; McKim, Karen L; Gollapudi, B Bhaskar; Moore, Nigel P; Haber, Lynne T; Allen, Bruce; Moore, Martha M 2017. Dose and temporal evaluation of ethylene oxide-induced mutagenicity in the lungs of male big blue mice following inhalation exposure to carcinogenic concentrations. *Environmental and molecular mutagenesis*, 58 3, 122-134.

Foundation for Chemistry Research and Initiatives, Electric Power Research Institute, American Chemistry Council, University of Minnesota, Aluminum Company of America, Dupont Company, American Journal of Epidemiology, Materion, Ramboll Environ Corporation, Hellenic Center for Disease Control and Prevention, Department of Justice Canada, Exponent, and Monsanto Corporation.

Dr. Checkoway recently received funding from the American Chemistry Council to review epidemiology literature on formaldehyde and non-Hodgkin's lymphoma (2017) and from the Monsanto Corporation to review epidemiological literature on polychlorinated biphenyls and non-Hodgkin's lymphoma (2016-19). These projects do not create a conflict interest and will not interfere with Dr. Checkoway's ability to provide an impartial, technically sound, and objective review of the subject matter.

Kenny Crump

Kenny S. Crump is retired and works as a private consultant. He has a Ph. D. in mathematics from Montana State University. Dr. Crump was a Professor of mathematics at Louisiana Tech from 1966 through 1980. Later he held positions in several companies, including K S Crump and Company, Inc. and Environ International Corporation. Dr. Crump's research has primarily involved development and application of statistical methodologies for quantitative assessment of health risks, primarily cancer, from exposures to toxic substances. The U. S. EPA has for many years used a statistical model developed by Dr. Crump to quantify cancer risk from toxic chemicals. Dr. Crump also initiated the "benchmark approach" for setting exposure standards for toxic chemicals that is now widely used in the U.S. and abroad. Dr. Crump is an author of more than 150 peer-reviewed scientific publications and book chapters. Many of these deal with risks to human health from specific substances, notably dioxin, asbestos, and diesel exhaust.

Dr. Crump has served as an advisor and on advisory committees for various bodies, including the USEPA's Science Advisory Board and Science Advisory Panel, the National Center for Toxicological Research's Science Advisory Board, the Mickey Leland National Urban Air Toxics Research Center's Science Advisory Panel, the National Institute of Environmental Health Sciences Board of Scientific Counselors, and the National Toxicology Program Board of Scientific Counselors. He has served on six National Academies of Science Committees and has been an official advisor to the World Health Organization, Health Canada, and the Province of Ontario. Dr. Crump is an elected Fellow of the American Statistical Association and the Society for Risk Analysis, and has received distinguished achievement awards from both of these organizations. During the past five years Dr. Crump has received funding from the Engine Manufacturers Association, the Canadian National Railroad, Tox-Strategies, and the U.S. EPA (Science Review Board on glyphosate).

Dr. Crump is not currently working on ethylene oxide; however, in 1983 he reviewed an OSHA quantitative risk assessment on EtO and presented his independent findings in an OSHA hearing. He has stated that his previous work on ethylene oxide was over 30 years ago and will not interfere with his objective evaluation of the ethylene oxide assessment. For the U.S. EPA, Dr. Crump served on a Science Review Board on glyphosate. This activity does not create a conflict interest and will not interfere with Dr. Crump's ability to provide an impartial, technically sound, and objective review of the subject matter.

Bette Meek

M.E. (Bette) Meek is the Associate Director of Chemical Risk Assessment at the McLaughlin Centre for Risk Science, Faculty of Medicine, University of Ottawa. Previously she contributed to and managed

several chemical risk assessment programs in Health Canada, most recently the assessment of Existing Substances under the Canadian Environmental Protection Act (CEPA). Dr. Meek received her Ph.D. in risk assessment from the University of Utrecht, the Netherlands and earned a M.Sc. in Toxicology from the University of Surrey, U.K. Dr. Meek has contributed to or led initiatives in developing methodology in chemical risk assessment, including mode of action, chemical specific adjustment factors, physiologically-based pharmacokinetic modeling, combined exposures, and predictive modeling. These initiatives have involved collaborations with a range of international organizations and national Agencies. She has co-authored approximately 200 peer-reviewed publications and more than 75 government publications. She has served on panels and/or as an advisor for national and international agencies and organizations, including Health Canada, U.S. EPA, U.S. National Academy of Sciences, World Health Organization, International Labour Organization, the European Joint Research Centre and the Agency for Food, Environmental and Occupational Health and Safety of France (ANSES). Dr. Meek received the Arnold J. Lehman Award from the Society of Toxicology. She has twice received the highest award of excellence of the Public Service of Canada, and several Health Canada Excellence in Science Awards. During the past five years, Dr. Meek has received external funding from the World Health Organization and Health Canada.

Dr. Meek contributed to the 2001 Health Canada assessment on ethylene oxide while managing the Priority Substances Program at Health Canada and was a contributing author on a resulting published manuscript and a 2003 World Health Organization Concise International Chemical Assessment Document. Dr. Meek has stated that the Health Canada EtO assessment was conducted in 2001 and that the relevant database and methodology has likely moved on significantly since then; therefore, she is confident that she can objectively review and comment on the TCEQ assessment. Dr. Meek has no conflicts of interest for this peer review.

David Savitz

David A. Savitz is Professor of Epidemiology and Interim Chair of the Department of Epidemiology in the School of Public Health at Brown University, with joint appointments in Obstetrics and Gynecology and Pediatrics in the Alpert Medical School. Previously, he was a professor at Mount Sinai School of Medicine and at the University of North Carolina, and an Assistant Professor at the University of Colorado School of Medicine. Dr. Savitz earned a Ph.D. in Epidemiology from the University of Pittsburgh Graduate School of Public Health. His epidemiological research has addressed environmental hazards in the workplace and community, reproductive health outcomes, and environmental influences on cancer. He has done extensive work on health effects of nonionizing radiation, pesticides, drinking water treatment by-products, and perfluorinated compounds. He is the author of nearly 350 papers in professional journals and editor or author of three books. He was President of the Society for Epidemiologic Research and the Society for Pediatric and Perinatal Epidemiologic Research. Dr. Savitz is an elected member of the National Academy of Medicine (NAM) where he has served on more than a dozen committees, five of which he chaired. He currently chairs the Health Effects Institute Research Committee and the NAM Committee to Review the Long-Term Effects of Antimalarial Drugs. From 2013-2017 he served as Vice President for Research at Brown University. During the past five years, Dr. Savitz has received research grants from the National Institutes of Health, the Department of Defense, and the Agency for Toxic Substances and Disease Registry (subcontract with RTI). He has also served as a consultant on several legal cases. None of these involved ethylene oxide.

Dr. Savitz has collaborated professionally with Dr. Kyle Steenland on several projects and publications, but not on ethylene oxide. Dr. Savitz has stated that his professional association with Dr. Steenland will

not interfere with his objective evaluation of the ethylene oxide assessment. Dr. Savitz has no conflicts of interest for this peer review.

Rita Schoeny

Rita Schoeny retired from the U.S. EPA in 2015 after 30 years and is currently serving as a consultant in risk assessment and science policy. Her most recent positions at U.S. EPA were as Senior Science Advisor for the Office of Science Policy, Office of Research and Development; and as the Director of the Risk Assessment Forum in EPA's Office of the Science Advisor. Dr. Schoeny received her Ph.D. in microbiology from the University of Cincinnati. She regularly lectures at colleges and universities, and has given training and lectures on risk assessment, science policy and toxicology in many areas of the world. She has been responsible for major assessments and programs in support of legislative mandates including the Safe Drinking Water Act, Clean Water Act, Clean Air Act, and Food Quality and Protection Act. Dr. Schoeny has published in the areas of toxicity of PCBs, PAHs, mercury, and drinking water contaminants (including disinfectant byproducts and microbes); assessment of complex environmental mixtures; and principles and practice of human health risk assessment. Recent work includes frameworks for human health risk assessment, interpretation of DNA adduct data for risk assessment, evaluation of episodic and less-than-lifetime exposure to carcinogens, new approaches to dose response assessment (including application of benchmark and other modeling procedures to chemicals including carcinogens), OECD guidelines for genetic toxicity testing, quantitative approaches to genetic toxicology, Adverse Outcome Pathways and Mode of Action, and approaches to cancer risk assessment. She has served on WHO committees and a National Academy of Sciences committee on risk assessment of complex mixtures. Dr. Schoeny is the recipient of numerous awards including EPA's Science Achievement Award for Health Sciences and the FDA Teamwork Award for publication of national advice on mercury-contaminated fish. She is an elected Fellow of the Society for Risk Analysis.

During the past five years, Dr. Schoeny has consulted on a paid basis with the American Chemistry Council Adverse Outcome Pathways project, Procultivos ANDI, the University of Cincinnati RSC, and Health Canada (subcontract with University of Cincinnati). She also does unpaid work for Save EPA, Environmental Protection Network, and other groups. This unpaid work has included contributing to discussions of current U.S. EPA published assessments in support of TSCA, as well as reviewing comments made by several retired U.S. EPA scientists and managers.

Dr. Schoeny worked for 30 years for the U.S. EPA, but she did not work on ethylene oxide. She has done some consulting work for the American Chemistry Council on Adverse Outcomes Pathways, but none of that work has been on ethylene oxide. None of her unpaid work has involved ethylene oxide, and Dr. Schoeny notes that her affiliation with these groups does not influence her scientific opinion on ethylene oxide or other chemicals. Dr. Schoeny has recently shared personal opinions about the U.S. EPA in a public blog. Regarding this, she notes: "A few years ago I contributed to a blog of the Union of Concerned Scientists. In this post I called into question the stated intention of the current administration to overturn or not enforce regulations promulgated by the U.S. EPA. I noted the lack of process in this endeavor by contrast to the established, iterative process for review, comment, and stakeholder involvement in issuing environmental regulations as well as the scientific evaluations in support of risk management choices. I state that as scientists we need to 'insist on the validity and thoroughness of our discipline'. I continue to uphold this opinion, and I here note that as part of this stance I have offered critiques of science based solely on the data available and my best scientific judgement. I note that my publications reflect this." Dr. Schoeny has stated that her previous employment by the U.S. EPA and consulting work with the American Chemistry Council and others will not bias her review and will not

interfere with her objective evaluation of the ethylene oxide assessment. Dr. Schoeny has no conflicts of interest for this peer review.

Risk Science Center

The UC RSC evaluated the potential for conflict of interest for RSC as an organization and for RSC personnel. The following activities related to ethylene oxide and relationships with the interested parties were identified.

- When she was an employee of Toxicology Excellence for Risk Assessment (TERA), Dr. Lynne Haber of the RSC was the lead of a project that investigated the mode of action for ethylene oxide-induced lung tumors in male mice. The project was conducted through a Cooperative Research and Development Agreement (Grant Number: CRADA E7229.11) between the National Center for Toxicological Research (NCTR) of the U.S. Food and Drug Administration and TERA. Funding for TERA and support for NCTR under the CRADA was provided by Ethylene Oxide and Derivatives Producers' Association and Lower Olefins Sector Group of the European Chemical Industry Association (CEFIC). This funding paid for materials at NCTR and Dr. Haber and other's time to work with NCTR on study design, interpretation of results, and preparation of publications. Dr. Haber coauthored two papers with the project team.
- In a second project while with TERA, Dr. Haber worked with Environ, with funding from CEFIC, to use the modified Hill criteria to integrate all of the MOA data, including the data obtained by NCTR, for an overall MOA evaluation. The focus of this work was MOA for lung cancer in mice.

These two projects were completed by 2014 while Dr. Haber was employed by TERA, and the RSC has had no work on ethylene oxide. Because Dr. Haber's projects are completed and no other work is planned, the RSC and Dr. Haber have no financial conflict of interest with this peer review work order. The MOA of EtO-induced mouse lung tumors is not an issue or outstanding question for the current ethylene oxide assessment, and the work Dr. Haber did is not directly relevant to the human cancer dose-response assessment that is the subject of this peer review. Therefore, this prior work does not create a situation that would cause Dr. Haber (or the RSC) to have a bias regarding the ethylene oxide DSD.

- Dr. Michael Dourson previously worked for the RSC and was a member of the EPA Science Advisory Board (SAB) for several years. He may have been involved in discussions of EPA's ethylene oxide work as an SAB member. Dr. Dourson left the RSC in 2017.
- A recent publication on ethylene oxide that is cited in the TCEQ DSD was co-authored by two individuals who previously worked for the RSC (Melissa Vincent and Andy Maier). None of the work for that publication took place while they were employed by the RSC. The RSC collaborates with Ms. Vincent and Dr. Maier on other projects unrelated to ethylene oxide. The RSC does not believe that these collaborations will cause the RSC staff to be biased regarding the ethylene oxide DSD or in conducting this peer review.

The RSC worked on two peer review projects in the last five years related to ethylene glycols (which are made from ethylene oxide).

- For the American Chemistry Council Ethylene Oxide/Ethylene Glycols Panel (which provided comments on the proposed DSD), Ms. Patterson of the RSC organized a panel of experts to perform a quality assurance review of two publications that derived risk values for diethylene glycol (DEG) and ethylene glycol (EG) for inclusion on the International Toxicity Estimates for Risk (ITER) database. This review project was completed in April 2019 with loading of the summaries on the ITER database.

- For Health Canada, the RSC (under a contract held by Toxicology Excellence for Risk Assessment) organized a letter peer review of a screening assessment on ethylene glycol ethers. The review was completed in June 2016.

These review projects did not involve ethylene oxide and involved peer review. No other work on ethylene glycols is planned.

Appendix D –References Identified by Experts

- Allen, B., Zeiger, E., Lawrence, G., Friedman, M., and Shipp, A. 2005. Dose-response modeling of in vivo genotoxicity data for use in risk assessment: some approaches illustrated by an analysis of acrylamide. *Reg. Tox. Pharm.* 41:6-27.
- Allen, B., Vincent, M., Liska, D., and Haber, L. 2016. Meta-regression analysis of the effect of *trans* fatty acids on low-density lipoprotein cholesterol. *Food and Chemical Toxicology* 98(Pt B):295-307. DOI: 10.1016/j.fct.2016.10.014.
- Allen, B., Shao, K., Hobbie, K., Mendez, W., Lee, J., Cote, I., Druwe, I., Gift, J., Davis, J. 2019. Bayesian Hierarchical Meta-Regression of Epidemiological Studies, Part 2: Dose-Response Modeling and Target Population Predictions. Submitted.
- Banda, M., K.L. McKim, L.T. Haber, J.A. MacGregor, B.B. Gollapudi, and B.L. Parsons. 2015. Quantification of *Kras* mutant fraction in the lung DNA of mice exposed to aerosolized particulate vanadium pentoxide by inhalation. *Mutat Res Genet Toxicol Environ Mutagen.* 789-790:53-60.
- Bisanti, L., Maggini, M., Raschetti, R., Alegiani, S. S., Ippolito, F. M., Caffari, B., . . . Ponti, A. 1993. Cancer mortality in ethylene oxide workers. *British Journal of Industrial Medicine*, 50(4), 317-324. doi:10.1136/oem.50.4.317
- Bulka, C., Nastoupil, L. J., Koff, J. L., Bernal-Mizrachi, L., Ward, K. C., Williams, J. N., . . . Flowers, C. R. 2016. Relations between residential proximity to EPA-designated toxic release sites and diffuse large B-cell lymphoma incidence. *Southern Medical Journal*, 109(10), 606-614. doi:10.14423/SMJ.0000000000000545
- Coggon, D., Harris, E. C., Poole, J., & Palmer, K. T. 2004. Mortality of workers exposed to ethylene oxide: Extended follow up of a british cohort. *Occupational and Environmental Medicine*, 61(4), 358-362. doi:10.1136/oem.2003.008268
- Greenberg, H. L., Ott, M. G., & Shore, R. E. 1990. Men assigned to ethylene oxide production or other ethylene oxide related chemical manufacturing: A mortality study. *British Journal of Industrial Medicine*, 47(4), 221-230. doi:10.1136/oem.47.4.221
- Hagmar, L., Mikoczy, Z., & Welinder, H. 1995. Cancer incidence in Swedish sterilant workers exposed to ethylene oxide. *Occupational and Environmental Medicine*, 52(3), 154-156. doi:10.1136/oem.52.3.154
- Hogstedt, C., Rohlén, O., Berndtsson, B. S., Axelson, O., & Ehrenberg, L. 1979. A cohort study of mortality and cancer incidence in ethylene oxide production workers. *British Journal of Industrial Medicine*, 36(4), 276-280. doi:10.1136/oem.36.4.276
- Kardos, L., Széles, G., Gombkötő, G., Szeremi, M., Tompa, A., & Ádány, R. 2003. Cancer deaths among hospital staff potentially exposed to ethylene oxide: An epidemiological analysis: Ethylene oxide exposure and cancer deaths. *Environmental and Molecular Mutagenesis*, 42(1), 59-60. doi:10.1002/em.10167

- Kiesselbach, N., Ulm, K., Lange, H. J., & Korallus, U. 1990. A multicentre mortality study of workers exposed to ethylene oxide. *British Journal of Industrial Medicine*, 47(3), 182-188. doi:10.1136/oem.47.3.182
- Kiran, S., Cocco, P., Mannetje, A., Satta, G., D'Andrea, I., Becker, N., . . . Boffetta, P. 2010. Occupational exposure to ethylene oxide and risk of lymphoma. *Epidemiology*, 21(6), 905-910. doi:10.1097/EDE.0b013e3181f4cc0f
- Manjanatha, M.G., S. D. Shelton, L.T. Haber, B. Gollapudi, J.A. MacGregor, and M.M. Moore. 2015. Evaluation of *cII* mutations in lung of male Big Blue mice exposed to vanadium pentoxide by inhalation for up to 8 weeks. *Mutat Res Genet Toxicol Environ Mutagen*. 789-790:46-52.
- Manjanatha, M., Shelton, S., Chen, Y., Parsons, B., Myers, M., Gollapudi, B., Moore, N., Haber, L., Allen, B., and Moore, M. 2017. Dose and temporal evaluation of ethylene oxide-induced mutagenicity in the lungs of male big blue mice following inhalation exposure to carcinogenic concentrations. *Environmental and Molecular Mutagenesis*, 58(3), pp.122-134.
- Mikoczy, Z., Tinnerberg, H., Björk, J., Albin, M., EpiHealth: Epidemiology for Health, Lund University, . . . Lunds universitet. 2011. Cancer incidence and mortality in Swedish sterilant workers exposed to ethylene oxide: Updated cohort study findings 1972-2006. *International Journal of Environmental Research and Public Health*, 8(6), 2009-2019. doi:10.3390/ijerph8062009
- Moore, M., Heflich, R., Haber, L., Allen, B., Shipp, A., and Kodell, R. 2008. Analysis of in vivo mutation data can inform cancer risk assessment. *Regulatory Toxicology and Pharmacology* 51:151.
- Norman, S. A., Norman, S. A., Berlin, J. A., Berlin, J. A., Soper, K. A., Soper, K. A., . . . Stolley, P. D. 1995. Cancer incidence in a group of workers potentially exposed to ethylene oxide. *International Journal of Epidemiology*, 24(2), 276-284. doi:10.1093/ije/24.2.276
- Olsen, G. W., Lacy, S. E., Bodner, K. M., Chau, M., Arceneaux, T. G., Cartmill, J. B., . . . Boswell, J. M. 1997. Mortality from pancreatic and lymphopietic cancer among workers in ethylene and propylene chlorohydrin production. *Occupational and Environmental Medicine*, 54(8), 592-598. doi:10.1136/oem.54.8.592
- Parsons BL, Manjanatha MG, Myers MB, McKim KL, Shelton SD, Wang Y, Gollapudi BB, Moore NP, Haber LT, Moore MM. 2013. Temporal changes in K-ras mutant fraction in lung tissue of big blue B6C3F₁ mice exposed to ethylene oxide. *Toxicol Sci*. 2013 Nov; 136(1):26-38. doi: 10.1093/toxsci/kft190
- Steenland, K., Stayner, L., Greife, A., Halperin, W., Hayes, R., Hornung, R., & Nowlin, S. 1991. Mortality among workers exposed to ethylene oxide. *The New England Journal of Medicine*, 324(20), 1402-1407. doi:10.1056/NEJM199105163242004
- Teta, M. J., Benson, L. O., & Vitale, J. N. 1993. Mortality study of ethylene oxide workers in chemical manufacturing: A 10 year update. *British Journal of Industrial Medicine*, 50(8), 704-709. doi:10.1136/oem.50.8.704

Vincent, M. J., Allen, B., Palacios, O. M., Haber, L. T., & Maki, K. C. 2019. Meta-regression analysis of the effects of dietary cholesterol intake on LDL and HDL cholesterol. *The American journal of clinical nutrition*, 109(1), 7-16.

Ward, E. M., Fajen, J. M., Ruder, A. M., Rinsky, R. A., Halperin, W. E., & Fessler-Flesch, C. A. 1995. Mortality study of workers in 1,3-butadiene production units identified from a chemical workers cohort. *Environmental Health Perspectives*, 103(6), 598-603. doi:10.1289/ehp.95103598

Wong, O., & Trent, L. S. 1993. An epidemiological study of workers potentially exposed to ethylene oxide. *British Journal of Industrial Medicine*, 50(4), 308-316. doi:10.1136/oem.50.4.308