Response to External Peer Review Comments Received on the Ethylene Oxide Development Support Document

CAS Registry Number: 75-21-8

Response to Comments

May 15, 2020

TEXAS COMMISSION ON ENVIRONMENTAL QUALITY

Table of Contents

Ethylene Oxide DSD External Peer Review Response to Comments

Below are the Expert Peer Reviewers comments, and the TCEQ's responses, to the charge questions provided by TCEQ for the review of the Ethylene Oxide (EtO) Carcinogenic Dose-Response Assessment Development Support Document (DSD) and the Response to Public Comments on the first draft of the DSD. Where the Expert had multiple comments on a single charge question, the TCEQ has provided separate responses to these comments for ease of reading and understanding of the TCEQ's response. In the interest of brevity and conciseness, some peer reviewer comments have not been replicated in their entirety below, and in some cases shorter excerpts have been included that convey the substantive content of the comment. In some cases, some of the text within a comment has been rearranged to fit into topic categories for ease of response. In the case where an expert has no comments on a charge question, then no comments from that expert are included in the text below. The DSD sections referred to in the charge questions correspond to sections of the DSD that was revised after public comment (i.e., revised DSD dated January 31, 2020), and not to the final DSD.

Charge Question 1

The TCEQ conducted a systematic review of the literature relevant to the derivation of an inhalation unit risk factor for EtO (see DSD [draft dated 1/31/20] 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

Comment:

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 C). Of these, findings from the following article not reviewed would add valuable information to the TCEQ review. See response to Q.4 [charge question 4] for details

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

Response to Expert 1:

TCEQ appreciates the work that Expert 1 did to identify these studies. Examination of the provided reference list in the context of the systematic review indicates that these studies were captured in the TCEQ systematic review. The choices made for each of the studies are noted below and can be found in Appendix 1 of the DSD, which describes the details of the EtO systematic review. The systematic review was intended to inform a quantitative dose-response assessment specifically to derive a chronic carcinogenicity toxicity factor, and therefore the study inclusion was narrowed to those studies that could inform a dose-response assessment, and not all studies in the more broad literature of EtO and cancer in humans.

Bisanti et al. 1993 – Listed in Table 19 of Appendix 1, excluded due to lack of exposure or dose-response information required for deriving a toxicity factor.

Bulka et al. 2016 – This study was excluded at the SWIFT-Review stage because it did not contain the ethylene oxide MeSH term, and it was a proximity study that did not have exposure or dose-response information required for deriving a toxicity factor.

Coggon et al. 2004 – Listed in Table 19 of Appendix 1, excluded due to lack of exposure or dose-response information required for deriving a toxicity factor.

Greenberg et al. 1990 – Listed in Table 19 of Appendix 1, excluded because a more recent update of the study cohort was available.

Hagmar et al. 1995 – Listed in Table 19 of Appendix 1, excluded because a more recent update of the study cohort was available.

Hogstedt et al. 1979 – Listed in Table 19 of Appendix 1, excluded because a more recent update of the study cohort was available (this study is listed as Hogstedt et al. 1979a in Table 19).

Kardos et al. 2003 – Listed in Table 19 of Appendix 1, excluded due to lack of exposure or dose-response information required for deriving a toxicity factor.

Kiesselbach et al. 1990 – Listed in Table 19 of Appendix 1, excluded due to lack of exposure or dose-response information required for deriving a toxicity factor.

Kiran et al. 2010 – Listed in Table 19 of Appendix 1, excluded due to lack of exposure or dose-response information required for deriving a toxicity factor

Mikoczy et al. 2011 – Listed in Table 20 of Appendix 1, taken through the systematic review and evidence integration process.

Norman et al. 1995 – Listed in Table 19 of Appendix 1, excluded due to lack of exposure or dose-response information required for deriving a toxicity factor.

Olsen et al. 1997 – Listed in Table 19 of Appendix 1, excluded due to lack of exposure or dose-response information required for deriving a toxicity factor.

Steenland et al. 1991 – Listed in Table 19 of Appendix 1, excluded because a more recent update of the study cohort was available.

Teta et al. 1993 – Listed in Table 19 of Appendix 1, excluded because a more recent update of the study cohort was available.

Ward et al. 1995 – This study was excluded at the SWIFT-Review stage because it did not contain the ethylene oxide MeSH term, and in fact was a study about carcinogenic effects of 1,3-butadiene specifically in workplaces where ethylene oxide was not present.

Wong et al. 1993 – Listed in Table 19 of Appendix 1, excluded due to lack of exposure or dose-response information required for deriving a toxicity factor.

Expert 2

Comment 1:

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

Response to Expert 2 Comment 1:

Thank you for your feedback that the TCEQ EtO systematic review was done appropriately.

Comment 2:

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

Relevant endpoints examined	*- Study focused solely on cytotoxicity* *- Study only measured sister chromatid exchanges (SECs), protein adducts, or chromosomal changes*

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.

Response to Expert 2 Comment 2:

TCEQ agrees that if we conduct an independent assessment of the mode of action (MOA) then studies that were excluded from the current systematic review, including those that focus only on cytotoxicity or chromosomal aberrations, will have to be evaluated. In that case, we would conduct a separate systematic review to specifically address the question of MOA. However, we have decided at this time not to conduct an independent evaluation of EtO's carcinogenic MOA for this DSD (discussed more in subsequent responses).

Comment 3:

> *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.

Response to Expert 2 Comment 3:

This language in the systematic review description in Appendix 1 refers to the availability of both human and animal dose-response data to use for the derivation of a carcinogenic toxicity

factor. Because the TCEQ generally prefers human data over animal data for developing toxicity factors (TCEQ 2015), we derived the toxicity factor using the human data. Elsewhere in the DSD the discussion of the human data includes the weaknesses in the studies and the limitations in extrapolating the available human data, noting that the data are less than sufficient in providing information about the low-dose (i.e., ambient) region of the dose-response curve. However, that information is also not provided by the animal data, wherein the exposure concentrations are even higher than those in the available human data.

Comment 4:

> *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?

Response to Expert 2 Comment 4:

The reviewers were the TCEQ authors of the DSD. This information has been added to the DSD, Appendix 1 Section A1.4.

Comment 5:

> *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.]

Response to Expert 2 Comment 5:

The 70-year lifetime is a standard duration used by the TCEQ when conducting lifetable analyses for chronic ESL derivation (TCEQ 2015), as is noted in Chapter 2 and Appendix 1. The TCEQ thinks that the discussion of the 70-year versus 85-year duration in the Response to Public Comments document is thorough and that it is a good location for the comparison (as is noted by the expert peer reviewers in this document, who suggest that a lot of the comparisons of USEPA and TCEQ modeling choices are best located in the Response to Public Comments and/or in an appendix of the DSD, which the TCEQ has done).

Expert 3

Comment 1:

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.

Response to Expert 3 Comment 1:

Thank you for your feedback on the TCEQ EtO systematic review.

Comment 2:

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.

Response to Expert 3 Comment 2:

The results from the systematic review summarized in Appendix A1.5 demonstrate that after reviewing the data there are only 3 cohorts that can be considered for deriving a carcinogenic toxicity factor: NIOSH, UCC, and the Swedish Chemical Workers cohort. The Swedish cohort was not chosen because there was little quantitative exposure information and the cohort was relatively small (< 1,000 people, compared to > 2,000 people for UCC and > 17,000 people for NIOSH). A summary of the reasons for selection of the UCC and NIOSH cohorts has been added to the main text of the DSD in Section 4.1.2.

Expert 4

Comment:

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).

Response to Expert 4:

The TCEQ appreciates Expert 4's evaluation of the EtO literature for both the dose-response assessment and for the MOA analysis. The MOA analysis considerations are addressed in response to Expert 4's charge question 3 comments.

Expert 5

Comment:

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.

Response to Expert 5

The TCEQ appreciates Expert 5's evaluation of the EtO literature. The systematic review was conducted specifically for the EtO carcinogenic dose-response assessment and therefore was not designed to identify mechanistic studies that would be unsuitable for deriving a carcinogenic toxicity factor. The responses to charge question 3 address Expert 4's identified literature.

Expert 6

Comment:

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

Responses to Expert 6

Thank you for your feedback on the TCEQ EtO systematic review.

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 [draft dated 1/31/20] Sections 3.3 and 3.3.1). Section 3.3.1 of the DSD [draft dated 1/31/20] 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 2

Comment 1:

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.

> *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.

Response to Expert 2 Comment 1:

The TCEQ appreciates these comments on the EtO carcinogenic MOA evaluation completed by the USEPA and adopted and briefly summarized by the TCEQ. At this time the TCEQ is not going to undertake a more complete, independent evaluation of the EtO MOA for carcinogenesis, but the DSD has been substantially revised to reflect more thoroughly the MOA evidence and evaluation. Given that the weight of evidence is supportive of a mutagenic MOA for EtO, that there is little to no evidence for a non-mutagenic MOA, and that a health-protective model (i.e., a linear no-threshold model) is used for mutagenic carcinogens, the TCEQ has chosen to use the mutagenic MOA.

Comment 2:

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.

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.

> *"The evidence for causal associations between the key events and tumor formation has <u>strength</u> and <u>consistency</u>. 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.

> *"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?

"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. 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?

Response to Expert 2 Comment 2:

The TCEQ has revised the DSD to reflect the uncertainty in the determination of a mutagenic MOA for EtO, although at this time we will not be completing an independent comprehensive MOA evaluation of EtO. Because the weight of evidence is supportive of a mutagenic MOA for EtO and a protective dose-response model (i.e. a linear no-threshold model) is used for mutagenic carcinogens, the TCEQ has chosen to use the mutagenic MOA.

Comment 3:

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.

The point regarding study quality… should be made throughout.

Response to Expert 2 Comment 3:

The TCEQ has revised the DSD to make further note of study quality uncertainties for the MOA analysis.

Comment 4:

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.

Response to Expert 2 Comment 4:

The TCEQ has revised the DSD to modify this language and to provide more information about the uncertainty and study quality information for the MOA analysis.

Comment 5:

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.

> *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.

Response to Expert 2 Comment 5:

The TCEQ has revised the DSD to include discussion of clastogenicity and large DNA mutations in the MOA description. The TCEQ takes note of Expert 2's comments that clastogenicity and large DNA mutations are not strong support for low dose linearity and certainly not for supralinearity.

Comment 6:

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"

Response to Expert 2 Comment 6:

The TCEQ has revised the DSD to include this language.

Comment 7:

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.

Response to Expert 2 Comment 7:

The TCEQ has revised the DSD to add citations and to clarify that mutagenicity is a potential cause of carcinogenicity. In addition, the requirement for active processes to convert DNA damage to mutations is noted, as well as a reference to the Vincent et al. (2019) mechanism of action pathway that provides more details about potential steps in this process.

Expert 4

Comment 1:

I agree that the available data support a mutagenic mode of action.

Response to Expert 4 Comment 1:

The TCEQ appreciates Expert 4's feedback that the data support a mutagenic MOA for EtO carcinogenicity.

Comment 2:

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.

Response to Expert 4 Comment 2:

The TCEQ appreciates these comments on the EtO carcinogenic MOA evaluation completed by the USEPA and adopted by the TCEQ. At this time the TCEQ is not going to undertake a more complete, independent evaluation of the EtO MOA for carcinogenesis, but the DSD has been substantially revised to reflect more thoroughly the MOA evidence and evaluation. The TCEQ will carefully consider the peer reviewer's suggestions if and when an independent MOA evaluation for EtO carcinogenesis is conducted at a future point in time.

Expert 5

Comment:

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.

Response to Expert 5:

The TCEQ appreciates Expert 5's feedback that the data support a mutagenic MOA for EtO carcinogenicity.

Expert 6

Comment 1:

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.

Response to Expert 6 Comment 1:

The TCEQ appreciates Expert 6's feedback that the data support a mutagenic MOA for EtO carcinogenicity.

Comment 2:

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.)

Response to Expert 6 Comment 2:

The quote that "The DSD quite plainly states that as USEPA acknowledges, MOA data do not support their model." was, as the reviewer suggests, in reference to the USEPA's dose-response model. The TCEQ addresses the reviewer's concerns about supra-linearity in response to Charge Question 7.

Charge Question 3

The TCEQ adopts EPA's MOA analysis (DSD [draft dated 1/31/20] Section 3.3.1) and considers MOA as information relevant to the likely or expected shape of the dose-response (DSD [draft dated 1/31/20] 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 [draft dated 1/31/20] Section 3.4.1.4.2). Are the TCEQ conclusions concerning implications of the MOA scientifically defensible?

Expert 2

Comment 1:

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.

Response to Expert 2 Comment 1:

The TCEQ appreciates Expert 2's feedback on the default low-dose extrapolation policy decisions associated with an inadequate MOA demonstration. This feedback is consistent with the TCEQ's choice to use a Cox proportional hazards model for the EtO carcinogenic dose-response assessment, because that model is not only indistinguishable from linear across the range of doses modeled, including low doses, but also a point of departure is then used for linear low-dose extrapolation.

Comment 2:

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.

Response to Expert 2 Comment 2:

The TCEQ has included in the DSD a more thorough discussion of the potential early key events in EtO-carcinogenesis in humans, and notes in Chapter 4 and Appendix 1 that while this information is useful for the MOA determination, it is not robust enough to derive a chronic carcinogenicity toxicity factor.

Comment 3:

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.

Response to Expert 2 Comment 3:

The TCEQ has revised the DSD to include discussion of clastogenicity and large DNA mutations in the MOA description. The TCEQ takes note of Expert 2's comments that clastogenicity and large DNA mutations are not strong support for low dose linearity and certainly not for supralinearity.

Comment 4:

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

Response to Expert 2 Comment 4:

TCEQ appreciates Expert 2's feedback that MOA can and should be used for choosing low-dose extrapolation methods and for choosing dose-response models. This feedback is consistent with the TCEQ's choice to use a Cox proportional hazards model, which is informed by a

potentially linear low-dose MOA for EtO carcinogenicity. See Section 4.2.1 of the final DSD for further discussion.

Comment 5:

> *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).

Response to Expert 2 Comment 5:

The TCEQ thinks that the statement above about sublinearity is not restricted to the particular details of how the modeling characterizes the curve (e.g. whether the curve is continuous or is multi-linear) and so has not changed this language.

Comment 6:

> *3.4.1.2.2.2 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?

Response to Expert 2 Comment 6:

This statement has been changed in the DSD and no longer refers to male breast cancer.

Comment 7:

> *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.

Response to Expert 2 Comment 7:

The TCEQ has modified the DSD to further discuss the MOA data for EtO carcinogenesis and has included more explicit language about the shapes of dose-response curves that are supported by the MOA data (Section 4.2.1).

Expert 3

Comment:

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.

Response to Expert 3:

TCEQ appreciates Expert 3's feedback that MOA can and should be used for choosing low-dose extrapolation methods and dose-response models. This feedback is consistent with the TCEQ's choice to use a Cox proportional hazards model, which is informed by a potentially linear low-dose MOA for EtO carcinogenicity (see Section 4.2.1 of the final DSD for further discussion).

Expert 4

Comment 1:

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.

Response to Expert 4 Comment 1:

The TCEQ appreciates Expert 4's feedback on the importance of MOA for low-dose extrapolation, but that policy decisions often must still be made. It is important to note that even without the weight of evidence supporting a mutagenic MOA for EtO (or any-MOA relevant information for that matter), the same linear low-dose extrapolation method would be used as a conservative science policy choice, consistent with standard practice for regulatory agencies such as the TCEQ (e.g., TCEQ 2015).

Comment 2:

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.

Response to Expert 4 Comment 2:

The TCEQ has included in the DSD a more thorough discussion of the potential early key events in EtO-carcinogenesis in humans, and notes in Chapter 4 and Appendix 1 that while this information is useful for the MOA determination, it is not robust enough to derive a chronic carcinogenicity toxicity factor. Furthermore, in regard to early key events being investigated at doses that are too high to inform the shape of the dose-response curve at much lower environmental levels, TCEQ notes that: (1) since lymphoid cancer drives the EtO carcinogenic assessment, perhaps the most relevant mutagenicity data is that in the bone marrow of mice exposed to 25-200 ppm EtO by inhalation *in vivo* (Recio et al. 2004, see Figure 1 below); (2) the overall linear dose-response for mutagenicity in bone marrow is consistent with a linear dose-response (see C-17 of USEPA 2016) and did not plateau even at exposure concentrations as high as 200 ppm; (3) studies show that EtO absorption and tissue concentrations are linearly related to inhalation EtO concentration, at least in the range of exposures used in the relevant studies (USEPA 2016); and (4) tissue concentrations of EtO are expected to be approximately equal in mice, rats, and humans exposed to a particular air concentration of EtO (USEPA 2016). These data are supportive of an overall linear dose-response across the entire dose range and not supportive of a more than linear (i.e., supra-linear) dose-response.

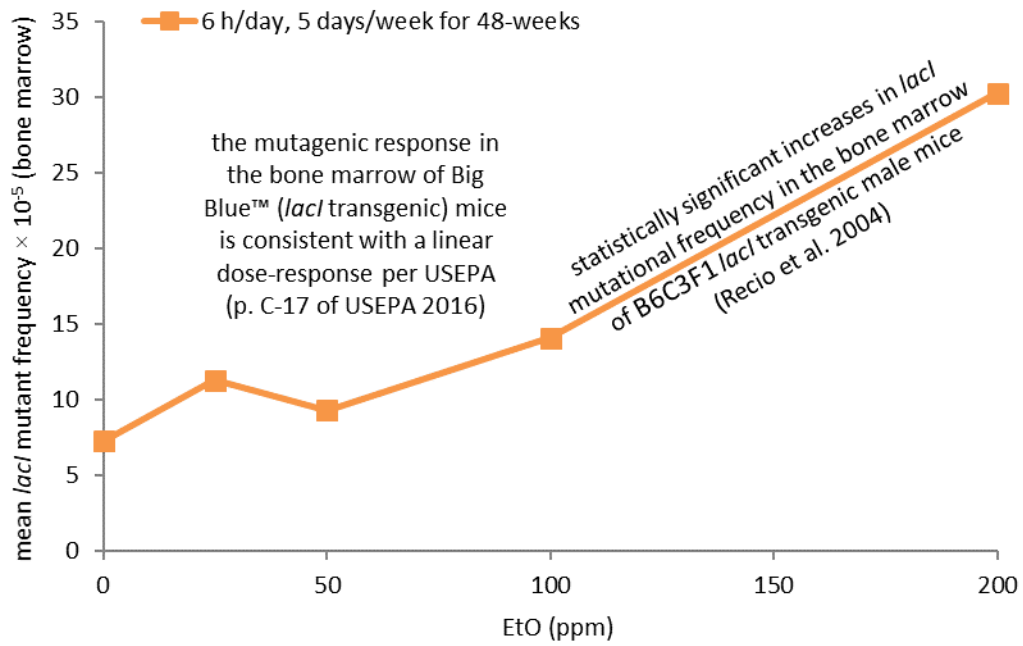

Figure 1. Overall linear dose-response for EtO-induced mutations in the bone marrow of Big Blue™ mice (based on data from Recio et al. 2004; Figure 3 in the final DSD).

Comment 3:

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).

Response to Expert 4 Comment 3:

In response to this and similar comments, the TCEQ has been more explicit in the final DSD concerning those decisions and assumptions that are made as a result of policy decisions (often informed by the TCEQ guidelines document, TCEQ 2015), and those that are based on the scientific data, as well as those that are a combination of the two. In the case of the EtO MOA, the TCEQ has added further discussion to the DSD and has determined that the MOA is likely to be direct mutagenicity. However, we also note that the TCEQ guidelines (2015) stipulate that either a mutagenic or an unknown MOA direct a non-threshold evaluation, which as standard practice consists of linear low-dose extrapolation. The Cox proportional hazards model is preferred under TCEQ guidelines for determining a point of departure for linear low-dose extrapolation.

Comment 4:

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).

Response to Expert 4 Comment 4:

The TCEQ responds to comments about endogenous formation and ground-truthing in response to Expert 4’s comments on questions 7 and 8, respectively.

Comment 5 (From Response to Charge Question 1):

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).

Response to Expert 4 Comment 5:

The TCEQ has included in the DSD a more thorough discussion of the potential early key events in EtO-carcinogenesis in humans, and notes in Chapter 4 and Appendix 1 that while this information is useful for the MOA determination, it is not robust enough to derive a chronic carcinogenicity toxicity factor. As mentioned above, in regard to key events such as mutagenicity the TCEQ notes that since lymphoid cancer drives the EtO carcinogenic assessment, perhaps the most relevant mutagenicity data is that in the bone marrow of mice exposed to 25-200 ppm EtO by inhalation in vivo (Recio et al. 2004, Figure 1). These data are consistent with a linear dose-response (see C-17 of USEPA 2016) and did not plateau even at exposure concentrations as high as 200 ppm. Furthermore, studies show that EtO absorption and tissue concentrations are linearly related to inhalation EtO concentration, at least in the range of exposures used in the relevant studies, and tissue concentrations of EtO are expected to be approximately equal in mice, rats, and humans exposed to a particular air concentration of EtO (USEPA 2016). These data are supportive of an overall linear dose-response across the entire dose range and not supportive of a more than linear (i.e., supra-linear) dose-response.

Expert 5

Comment 1:

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.

Response to Expert 5 Comment 1:

The TCEQ has edited the DSD to include the definition of “supra-linear” as a model with a dose-response curve that is steeper than linear as illustrated in Figure 2 of the final DSD where the low-dose slope is steep beginning at zero dose and then transitions at higher doses to a shallower slope (as is noted by Expert 5). By TCEQ’s definition this can include curvilinear models or multi-part linear spline models with this same shape. The TCEQ now largely discusses the USEPA (2016) dose-response model as a linear two-piece spline, with a steeper slope in the

low-dose region. Much of the discussion of the USEPA's model has been removed from the main text of the DSD, as has the quote from Crump and Allen (1985).

Comment 2:

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.

Response to Expert 5 Comment 2:

The TCEQ is not familiar with literature that supports a receptor-type model for direct acting carcinogens that is based on "occupation" by mutation of sites in tumor suppressor genes or oncogenes. This hypothesis assumes a finite number of mutable sites that can contribute to carcinogenesis, and while there are not an unlimited number of genes that will impact carcinogenesis, there are a substantial number of genes, likely more than we know (https://www.ncbi.nlm.nih.gov/books/NBK21662/). In addition, unlike with receptor binding that tends to be binary (i.e. the receptor is either "on" or "off"), genetic changes in carcinogenesis are nuanced with different mutations affecting the same gene in different ways that can change gene expression (increased or decreased), protein shape (affecting function, stability, degradation), protein targeting, binding to proteins, DNA, or RNA, etc. Therefore, because of the absence of literature and the nuanced considerations of gene mutations, the TCEQ does not find support for the hypothesis of gene mutations in carcinogenesis acting like a receptor occupation model. TCEQ has removed the Michaelis-Menton kinetics example from the final DSD because, while it is an example of a model with a mechanistic basis like the Chemical Industry Institute of Toxicology formaldehyde model cited as an example, the TCEQ did not mean to suggest that this type of model may apply to a carcinogen with a direct acting mutagenic MOA (where no metabolic activation is required). This type of model is not an appropriate comparator for the dose-response relationship of a direct acting carcinogen to human cancer mortality.

Comment 3:

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. 57[1] 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.

Response to Expert 5 Comment 3:

The TCEQ has substantially revised the DSD and made the basis for the decisions about the shape and selection of the dose-response model more clear, which include considerations of MOA, model fit, model validation, and policy decisions based on the TCEQ guidelines (2015). These points are specifically addressed in Section 4.2 of the final DSD. Specific to the comment, model fit and accuracy were very important considerations for TCEQ, and such evaluations (particularly model accuracy) fully support TCEQ's selected model. In regard to Expert 5's comment that the USEPA spline models do not clearly violate the MOA assumed for EtO, the TCEQ notes that USEPA (2016) acknowledged to the SAB that the MOA information for EtO does not support their selected dose-response, stating "the EPA is not aware of a mechanistic explanation" (p. I-29 of USEPA 2016; also see pp. I-34 and 4-71). Similarly, the TCEQ is not aware of any MOA or mechanistic data for EtO that would support the dose-response represented by USEPA's linear two-piece spline model, unlike the MOA information that supports the Cox proportional hazard model.

Comment 4:

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.

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

[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.

low-dose slope be different (presumably less) than estimated by USEPA, by how much, and in what ranges of exposure?

Response to Expert 5 Comment 4:

The TCEQ appreciates Expert 5's identification of an important point: the question of whether low-dose extrapolation should occur based on observations from much higher doses (thousands or millions of times higher, in fact). Both USEPA and TCEQ have done exactly that, used a high dose dataset to estimate the low-dose linear part of the dose-response curve, so TCEQ does not seek to rebut the standard regulatory agency practice of linear low-dose extrapolation based on high dose data (whether epidemiological or from animal studies). The only difference is the model used to do it. The TCEQ has concluded, based on the likely mutagenic MOA of EtO, that there should be a low-dose linear extrapolation from high dose data, using the Cox proportional hazards model as a standard model for that purpose supported by multiple lines of evidence.

The TCEQ has substantially rewritten the DSD so that the main text is focused on the EtO data and the analyses, decisions, and justifications made by the TCEQ in deriving a unit risk factor for EtO. Because the TCEQ has conducted an independent analysis of the data, there is no need to justify a change from the USEPA model. There is still some discussion of the different choices made by USEPA versus TCEQ, but these are largely restricted to the Response to Public Comments document and Appendix 6 of the final DSD.

Comment 5:

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).

Response to Expert 5 Comment 5:

The TCEQ has included in the DSD a more thorough discussion of the potential early key events in EtO-carcinogenesis in humans, and notes in Chapter 4 and Appendix 1 that while this information is useful for the MOA determination, it is not robust enough to derive a chronic carcinogenicity toxicity factor. In regard to key events such as mutagenic activity, the TCEQ notes above that since lymphoid cancer drives the EtO carcinogenic assessment, perhaps the most relevant mutagenicity data is that in the bone marrow of mice exposed to 25-200 ppm EtO by inhalation in vivo (Recio et al. 2004). These data are consistent with a linear dose-response (see C-17 of USEPA 2016) and did not plateau even at exposure concentrations as high as 200 ppm. The consideration of these data is supportive of an overall linear dose-response across the entire dose range and not supportive of a more than linear (i.e., supra-linear) dose-response.

Comment 6:

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 search[2]).

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,

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

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,

$$background\ lifetime\ cancer\ probability\ = \int_0^\infty (a + b \cdot e) \cdot p(e)\ de$$

$$= a\ + b * E(e)$$

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

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

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

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).

Response to Expert 5 Comment 6:

The TCEQ appreciates Expert 5's set of calculations that address considerations of how endogenous levels of EtO could contribute to total population lymphoid cancers. We do have some concerns about these calculations, however.

First, the assumption that 90% of lymphoid cancers are unrelated to EtO. One can argue that all lymphoid cancers have something to do with EtO because all humans have some levels of endogenous EtO. The unit risk attributable to EtO applies to how much cancer would be expected given a certain dose of EtO and is not restricted by hypothesized numbers of maximal cases. As is noted by Expert 5, in the case that all lymphoid cancers are related to EtO, then the resultant slope would be 0.013/ppb, or ten times higher than the back-of-the-envelope calculation, and 50 times higher than USEPA's calculation. The fact that the number of lymphoid cancers related to EtO is likely to be unknowable unfortunately means that this back-of-the-envelope calculation can be manipulated by changing the 90% to get the slope of interest.

In addition, if we continue with the expert's back-of-the-envelope calculations one could use the average exposure to EtO from the NIOSH data and determine if these slope estimates match what was observed in the NIOSH study. From Table 1 in Valdez-Flores (2010), the median occupational ppm-days in females (which is lower than males) is 1767 ppm-days (it is 2807 ppm-days in males). This is equivalent to an average environmental exposure concentration of 0.023 ppm from birth to 70 years (see footnote a of Table 11 in Section 4.3.2.1 for full conversion calculation). Multiplying 0.023 ppm by 1000 (to convert to ppb) and then by 0.0013/ppb results in an extra risk of 0.03. Then multiplying by the number of workers in the NIOSH study (17,493 in Table 2 of Valdez-Flores et al. 2010) the expected number of additional

lymphoid deaths would be 523. The actual "total" number of lymphoid deaths in the EtO-exposed population in the NIOSH study was 53; that is about 10 times lower than the 523 "additional" lymphoid deaths predicted by this calculation. Furthermore, the number of lymphoid deaths in the NIOSH study is less than expected using the population background rates (Table 32 in the DSD) so that the "additional" number of cases observed in the NIOSH study is zero; which is much smaller than the 523 additional cases estimated by this back-of-the-envelope calculations. This overestimation would be even greater if either: 1) USEPA's slope is used; or 2) the median exposure for males is used.

Comment 7:

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 B6C3F$_1$ 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.

Response to Expert 5 Comment 7:

The TCEQ appreciates Expert's 5 dose-response method suggestions, which include potentially using a lower slope in the truly low-dose area of the curve, with a transition to a higher slope at higher doses.

The TCEQ has reviewed the papers referenced by Expert 5 regarding the use of MOA or early key event data to quantitatively inform the low-dose area of the dose-response curve, with a focus on applying this to a carcinogenic toxicity factor derivation. While this method has

exciting possibilities for informing the shape of the dose-response curve at truly low doses (i.e. environmentally-relevant doses), the method does not seem to have been sufficiently developed to apply to toxicity factor derivation at this time. In addition, many of the examples provided in the papers were using datasets with well-developed exposure and outcome metrics (e.g., serum cholesterol levels and dietary cholesterol ingestion), whereas the available information for EtO exposure and non-cancer outcomes (e.g., sister chromatid exchanges) is far less reliable, nor could it be readily combined into a meta-regression analysis. The TCEQ will continue to investigate these methods for future use but at this time does not consider that these methods can be applied to this EtO carcinogenicity dose-response assessment.

Expert 6

Comment:

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).

Response to Expert 6:

Neither the draft nor final DSD use low-dose linearity as a basis to choose between the Cox proportional hazards model used by TCEQ and the spline model used by USEPA. The TCEQ has substantially revised the DSD and made the basis for the decisions about the shape of the dose-response model more clear, which include considerations of MOA, model fit, model validation, and policy decisions based on the TCEQ guidelines (2015). These points are specifically addressed in Section 4.2 of the final DSD.

Charge Question 4

The TCEQ conducted an evaluation of EtO's carcinogenic classification (DSD [draft dated 1/31/20] Section 3.3.2), and also evaluated breast cancer risk in humans as a potential cancer endpoint (DSD [draft dated 1/31/20] 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

Comment:

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.

Response to Expert 1:

The TCEQ appreciates Expert 1's review of the EtO-breast cancer epidemiological studies and their results. Following Expert 1's lead, the TCEQ evaluated the available human breast cancer results (summarized in Table 4 of the final DSD) and considered the question of internal and external referent analyses. The basis for using an internal referent is the healthy worker effect, which is a form of bias in epidemiology studies where, in theory, a population of workers may be healthier and less likely to develop the disease of interest compared to the general

population. Section 3.1.1.2 of the final DSD discusses the healthy worker effect and specifically addresses breast cancer. A recent very large occupational cohort study did not find evidence of a healthy worker effect for breast cancer (Kirkeleit et al. 2013). Therefore, the TCEQ considers that the external referent analyses provide valid information about the association between occupational exposure to EtO and breast cancer. Table 4 in Section 3.3.1.1.1.1 of the final DSD shows that the SIRs/SMRs for individual EtO studies (Steenland et al. 2003, Steenland et al. 2004, Norman et al. 1995, Coggon et al, 2004, and Mikoczy et al. 2011) are consistently not significantly elevated, with most actually being less than 1. The TCEQ also specifically reviewed the findings of the well-conducted studies (as suggested by Expert 1), which were identified in the meta-analysis of Vincent et al. (2019). That group found that a meta-analysis of the high-quality breast cancer studies (Steenland et al. 2003, Steenland et al. 2004, Mikoczy et al. 2011) generated a null association between EtO concentration and breast cancer (0.92, 95% CI 0.84, 1.02). This information, in addition to the considerations raised by Expert 3 in response to this charge question, altogether suggest that there is little evidence from human epidemiology studies of an association between occupational EtO exposure and human breast cancer (i.e., the overall weight of evidence is weak).

Expert 2

Comment 1:

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.

Response to Expert 2 Comment 1:

The TCEQ principally considered the breast cancer data in a WOE evaluation for the hazard assessment for EtO (Section 3.3.1.1). The TCEQ notes Expert 2's comment that the analyses as described in the DSD support the reasonable choice to not derive a URF using the breast cancer data.

Comment 2:

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.

Response to Expert 2 Comment 2:

In the final DSD, the TCEQ considered the human evidence as a whole in the determination that EtO is *likely to be carcinogenic to humans*, which included deciding (as IARC did) that the human data to support this determination are limited. The breast cancer results were considered in that "limited" determination. The TCEQ's hazard assessment of human breast cancer associated with EtO is now explicitly detailed in Section 3.3.1.1 of the DSD. The summary of that section is as follows:

In summary, the epidemiological evidence for EtO causing human breast cancer is very weak, with most of the available studies showing no association when the external reference population is used as a comparison group. This is the same conclusion reached by Marsh et al. (2019) in their recent meta-analysis, which found that there was no evidence from the epidemiology studies of a relationship between EtO exposure and breast cancer. The meta-analysis conducted by Vincent et al. (2019) reached a similar conclusion, stating that "Higher quality epidemiological studies demonstrated no increased risk of breast cancers". When considering the evidence from animal studies, the TCEQ found that while there was an increase in mammary tumors in mice chronically exposed to EtO (NTP, 1987), there was no increase in mammary tumors in rats chronically exposed to EtO (Snellings et al. 1984). In addition, IARC in 2019 released an assessment of tumor site concordance, which found that only 20% of the evaluated Group 1 chemicals showed site-concordance of mammary/breast tumors between animals and humans. While the MOA determination that EtO is carcinogenic through mutagenic and genotoxic mechanisms generically supports tumor sites at any location, there is no specific MOA or metabolic information that identifies breast tissues as a susceptible site. Therefore, the TCEQ determines that there is inadequate evidence for identifying breast cancer as a hazard of EtO exposure in humans.

Comment 3:

> *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.

Response to Expert 2 Comment 3:

The TCEQ has amended the DSD to discuss the WOE review of breast cancer in hazard assessment Section 3.3.1.1, which includes streamlining the text to remove judgmental writing and to clarify the analysis and determinations about the breast cancer data.

Comment 4:

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?

Response to Expert 2 Comment 4:

The TCEQ has deleted the referenced sentence from the DSD.

Comment 5:

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)?

Response to Expert 2 Comment 5:

TCEQ has clarified the discussion of the healthy worker effect in Section 3.1.1.2 of the DSD, as well as the discussion of Steenland et al.'s conclusion about under-ascertainment in Section 3.3.1.1.1.1, as follows:

"Steenland et al. (2003) stated that they used internal referents because of the potential for under-ascertainment; however since that study found that there was complete breast cancer ascertainment in the sub-cohort with interviews, the TCEQ still considers the external referent comparisons to be the most appropriate."

Comment 6:

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

Response to Expert 2 Comment 6:

The text of the DSD has been modified to reflect that the agents in the sentence are those agents studied by IARC in the assessment.

Comment 7:

Here [Appendix 6] 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.

Response to Expert 2 Comment 7:

The TCEQ principally considered the breast cancer data in a WOE evaluation for the hazard question: does EtO exposure cause human breast cancer? Although the TCEQ has ultimately chosen to designate EtO as a likely human carcinogen, this does not mean that all human cancers are caused by EtO. The primary types of cancer with some evidence in human studies are lymphoid cancers and breast cancer. As part of the process for deciding if there was likely to be a causal relationship between EtO and breast cancer in humans, the TCEQ investigated the various available data, which included animal studies showing possible increases in mammary tumors with EtO exposure. However, the IARC (2019) work demonstrates that the state of the science does not support tumor site concordance as a general principle. Thus, the positive mammary carcinogenesis results in one animal species (mice but not rats) do not provide support for a specific site of cancer in humans, and therefore does not inform the breast cancer

decision outside of providing information about the initial carcinogenesis determination. This information has now been clarified in the DSD.

Expert 3

Comment 1:

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.

Response to Expert 3 Comment 1:

While the Marsh et al. (2019) analysis did assess study quality, Expert 3 is correct that the meta-analysis was conducted on all the studies regardless of quality. In contrast, Vincent et al. (2019) did conduct a meta-analysis on EtO breast cancer cohort results that distinguished study quality, and they found that high or medium quality studies did not report an association between EtO and breast cancer, whereas the low quality study did show an association (Figure 1 of Vincent et al. 2019). The meta-analysis of the high quality studies generated an effect estimate of 0.92 (95% CI of 0.84, 1.02). This information has been clarified in the DSD.

Comment 2:

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 or 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.

Response to Expert 3 Comment 2:

The TCEQ appreciates the information provided by Expert 3 about the difficulties of interpreting the NIOSH cohort findings. We have updated the DSD to reflect this information.

As per Expert 3's suggestions, the TCEQ has carefully considered the information from both continuous and categorical analyses of the breast cancer data, as well as the potential for a healthy worker effect for breast cancer and external and internal referent comparisons. The information provided by the expert peer reviewers in response to this charge question are now considered in a separate section of the DSD that presents the WOE of a breast cancer hazard identification for EtO (Section 3.3.1.1.1). The final synthesis of this information is presented in response to Expert 2 Comment 2 and includes TCEQ's judgement call that there is insufficient evidence to consider EtO as a cause of human breast cancer. This is consistent with the Expert 3's conclusion, "I think that TCEQ made the right choice to not consider breast cancer in their quantitative assessment."

Expert 4

Comment 1:

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.

Response to Expert 4 Comment 1:

The TCEQ has added language to the DSD to clarify that Marsh et al. (2019) and Vincent et al. (2019) were re-analyses of available data, and not new data. TCEQ appreciates Expert 4's discussion about the Marsh and Vincent study findings and notes Expert 4's conclusion that there is limited to no consistent evidence of an association between EtO and breast cancer mortality. Notably, this is the case in workers exposed to EtO concentrations up to millions of times higher than environmental levels.

Comment 2:

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.

Response to Expert 4 Comment 2:

The TCEQ is principally considering whether the WOE demonstrates that EtO is causing breast cancer in humans. As part of the process for deciding if there was likely to be a causal relationship between EtO and breast cancer in humans, we investigated the various available data, which included animal studies showing possible increases in mammary tumors with EtO exposure. However, the IARC (2019) work demonstrates that the state of the science does not support tumor site concordance as a general principle. Thus, the positive mammary carcinogenesis results in one animal species (mice but not rats) do not provide support for a specific site of cancer in humans, and therefore does not inform the breast cancer decision outside of providing information about the initial carcinogenesis determination. This information has now been clarified in the DSD.

Comment 3:

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/dwq/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).

Response to Expert 4 Comment 3:

The TCEQ has revised the DSD to consider more logically and holistically the available epidemiology, animal, and MOA evidence for the determination of whether EtO causes human breast cancer. The TCEQ's hazard assessment of human breast cancer caused by EtO is now explicitly detailed in Section 3.3.1.1 of the final DSD. The summary of that section is as follows:

In summary, the epidemiological evidence for EtO causing human breast cancer is very weak, with most of the available studies showing no association when the external reference population is used as a comparison group. This is the same conclusion reached by Marsh et al. (2019) in their recent meta-analysis, which found that there was no evidence from the epidemiology studies of a relationship between EtO exposure and breast cancer. The meta-analysis conducted by Vincent et al. (2019) reached a similar conclusion, stating that "Higher quality epidemiological studies demonstrated no increased risk of breast cancers". When considering the evidence from animal studies, the TCEQ found that while there was an increase in mammary tumors in mice chronically exposed to EtO (NTP, 1987), there was no increase in mammary tumors in rats chronically exposed to EtO (Snellings et al. 1984). In addition, IARC in 2019 released an assessment of tumor site concordance, which found that only 20% of the evaluated Group 1 chemicals showed site-concordance of mammary/breast tumors between animals and humans. While the MOA determination that EtO is carcinogenic through mutagenic and genotoxic mechanisms generically supports tumor sites at any location, there is no specific MOA or metabolic information that identifies breast tissues as a susceptible site. Therefore, the TCEQ determined that there is inadequate evidence for identifying breast cancer as a hazard of EtO exposure in humans.

Expert 5 Comment:

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.

Response to Expert 5:

The TCEQ appreciates Expert 5's discussion about the breast cancer data and notes Expert 5's conclusion that the EtO dose-response assessment should be based on lymphohematopoietic cancers rather than breast cancer.

Expert 6

Comment:

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.

Response to Expert 6:

The TCEQ appreciates Expert 6's discussion about the breast cancer data and notes Expert 6's conclusion that the data is only suggestive that occupational exposures to EtO may increase the risk of breast cancer.

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

Comment:

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.

Response to Expert 1:

The TCEQ appreciates Expert 1's comment that mortality data may be a better surrogate for lymphoid cancer risk, which has a higher mortality rate, than for breast cancer risk.

Expert 2

Comment:

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.

Response to Expert 2:

After careful consideration of the comments received on this charge question, the TCEQ has decided not to try to address lymphoid cancer incidence through the available lymphoid cancer mortality data. The risk-based air concentrations and URFs are based on lymphoid cancer mortality. As discussed in TCEQ guidelines (TCEQ 2015), uncertainty is increased if the endpoint used in calculating excess risks (e.g., cancer incidence) is different than the endpoint used in the dose-response modeling (e.g., cancer mortality). It is most appropriate, when excess risks for the inference population are being calculated, for the health endpoint to be the same health endpoint as was used in the dose-response modeling. The computational details of the BEIR IV methodology are different for incidence and mortality (e.g., see Sielken and Valdez-Flores 2009). Accordingly, the TCEQ does not generally use a mortality-based exposure-response model as the basis for the calculation of excess risks for an incidence response (or vice versa). This DSD adheres to the general principle in TCEQ guidance (TCEQ 2015) that the health endpoint used for dose-response modeling and the excess risk calculation should match. Thus, since the available data are for mortality, lymphoid cancer mortality (not incidence) serves as the basis for the TCEQ's risk-based air concentrations and URFs.

Expert 3

Comment:

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.

Response to Expert 3:

TCEQ appreciates Expert 3's in-depth analysis of the important uncertainties of using cancer mortality versus incidence data. Section 4.3.2 of the DSD now includes a paragraph that summarizes TCEQ's decision to use the cancer mortality data to model cancer mortality, not incidence, and part of that decision is the consideration of uncertainty as expressed by Expert 3. As noted by Expert 3, TCEQ is making the use of the best available data when quantifying the

dose-response of cancer mortality, because incidence data are unavailable for the critical endpoint, lymphoid cancer.

Expert 4

Comment:

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.

Response to Expert 4:

The TCEQ appreciates Expert 4's concerns about the uncertainties of using mortality data to estimate cancer incidence. Section 4.3.2 of the DSD now includes a paragraph that summarizes TCEQ's decision to use the cancer mortality data to model cancer mortality, not incidence, and part of that decision is the consideration of uncertainty as expressed by Expert 4. The TCEQ does utilize the 95% UCL on lymphoid cancer mortality for additional conservatism, partly based on mortality versus incidence considerations.

Expert 5

Comment:

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).

Response to Expert 5:

The suggestion made by Expert 5 to derive risk estimates using lifetable analysis for incidence instead of mortality, while using the relative risks from mortality data, is indeed what USEPA did in their analysis. The TCEQ considered doing this, but ultimately decided not to because of the uncertainty in the assumptions of using a dose-response for mortality combined with incidence data. We agree that if this incidence-mortality hybrid approach is used, the survival rates need to be adjusted to reflect loss from follow-up (this method is described in Sielken and Valdez-Flores 2009). Section 4.3.2 of the DSD now includes a paragraph that summarizes TCEQ's decision to use the cancer mortality data to model cancer mortality, not incidence.

Expert 6

Comment:

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.

Response to Expert 6:

The TCEQ appreciates Expert 6's comments on the factors that affect the comparison of cancer incidence to mortality.

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 [draft dated 1/31/20] 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 2

Comment 1

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.

Response to Expert 2 Comment 1:

The TCEQ has substantially revised the DSD section on visual fit and has changed "corrected" to "recalculated" when referring to AIC for various models.

Comment 2

> *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.

Response to Expert 2 Comment 2:

The TCEQ has modified the DSD and the above statement is no longer present. Section 4.2 of the DSD makes clear the reasons for the choice of the Cox proportional hazards dose-response model.

Expert 4

Comment 1:

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).

Response to Expert 4 Comment 1:

The TCEQ appreciates Expert 4's comment that mechanistic basis is a primary consideration when making a model choice, with model fit criteria being a secondary consideration. The TCEQ implements this concept by using the MOA to decide on a model that is no more than linear overall (the Cox proportional hazards model is indistinguishable from linear across the range of doses modeled), which is also strongly supported by additional considerations (e.g., model fit and accuracy).

Comment 2:

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 ($\chi 2$ 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)

Response to Expert 4 Comment 2:

Section 4.2 of the final DSD addresses the parameters that the TCEQ used to consider dose-response model choice, including MOA information, AIC and p-value, and model accuracy when compared to the NIOSH cohort data that is the data source for the dose-response model. The TCEQ's NIOSH model is also validated using UCC data.

Comment 3:

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.

Response to Expert 4 Comment 3:

The TCEQ agrees with Expert 4 and uses individual data in the dose-response assessment, including for evaluation of model fit (e.g., AIC and p-values, visual fit).

Comment 4:

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.

Response to Expert 4 Comment 4:

The TCEQ has substantially revised the DSD and has clarified the critical focus on the choice of the dose-response model (formerly in Section 3.4, now in Section 4.2).

Comment 5:

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).

Response to Expert 4 Comment 5:

The TCEQ notes Expert 4's opinion that considerations of endogenous production does not add weight to the argument about model fit. This topic has been moved from the main text of the DSD to Appendix 5 where it is presented as biological context for EtO. The TCEQ does think that the accurate estimation of the NIOSH cohort findings is an important aspect of model choice that has been further discussed and clarified in section 4.2.3 of the DSD. The TCEQ has also clarified the MOA discussion in the final DSD, as noted in response to charge questions 2 and 3. In terms of additional "ground truthing", which expert 4 found informative, the accuracy of TCEQ's NIOSH model is now also validated using a different dataset (UCC data). Further discussion of ground truthing in response to Expert 4's suggestions is presented in response to charge question 8.

Expert 5

Comment 1:

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.

Response to Expert 5 Comment 1:

The TCEQ appreciates Expert 5's analysis of the method for calculating the AIC and p-values and notes their conclusion that the TCEQ has correctly calculated these parameters.

Comment 2:

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.

Response to Expert 5 Comment 2:

The TCEQ appreciates Expert 5's discussion of the difficulty of interpreting and making use of visual model fit, and the expert's suggestion about reliance on statistical diagnostics over visual fit. The x- and y-axis choices in the Visual Fit section (now in Appendix 6) have been more thoroughly described in the final DSD to clarify the purpose of changing the axis scales and to better explain the methodology used by the TCEQ to generate these graphs. The TCEQ only discusses visual fit (and only in an Appendix) because of USEPA's reliance on it. By contrast, the TCEQ does not rely on visual model fit as a primary consideration for model choice, but rather principally relies on MOA and various statistical diagnostics of model fit (i.e., AIC and p-values, statistical analyses of model accuracy), consistent with the comment.

Comment 3:

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.

Response to Expert 5 Comment 3:

The TCEQ has substantially rewritten the DSD so that the main text is focused on the EtO data and the analyses, decisions, and justifications made by the TCEQ in deriving a unit risk factor for EtO. Because the TCEQ has conducted an independent analysis of the data, there is no need to justify a change from the USEPA model. There is still some discussion of the different choices made by USEPA compared to TCEQ, but these are largely restricted to the Response to Public Comments document and Appendix 6 of the final DSD.

Section 4.2 of the final DSD addresses the parameters that the TCEQ used to consider dose-response model choice, including MOA information, AIC and p-value, and model accuracy when compared to the NIOSH cohort data that is the data source for the dose-response model. In contrast to the expert's suggestion that the linear two-piece spline and Cox proportional hazard models are essentially equal, and therefore health protectiveness should be the deciding factor, important factors do favor the Cox proportional hazards model. These include: 1) better alignment with the MOA for a direct-acting mutagenic carcinogen; 2) parsimony (i.e. fewer model components); and 3) accuracy in predicting lymphoid cancer in the primary NIOSH cohort dataset as well as in a validation analysis using the UCC cohort dataset. These facts show that the models are not equal as implied, but rather favor the use of the Cox proportional hazards model, which is why the TCEQ chose this model.

In addition, as discussed in Section 4.3.2.2, the URF derived by the TCEQ is quite health protective, because it assumes a statistically significant association between EtO dose and lymphoid cancer (while the modeling shows no difference from the null), with use of the 95% upper confidence limit providing further health protection. Furthermore, the URF is based on the more conservative estimate for lymphoid cancer in males (compared to males and females combined) and incorporates USEPA age-dependent adjustment factors, both of which result in a lower, more protective risk-based air concentration for the general public.

Expert 6

Comment 1:

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* [emphasis added with note from TCEQ: this second AIC should be 464.4], 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.

Response to Expert 6 Comment 1:

The TCEQ appreciates Expert 6's analysis of the method for calculating the AIC and p-values and notes that the TCEQ considered the determination of the knot in the spline model to be a parameter in the AIC determination, but did not include lag as a parameter because lag was not estimated based on the data. The TCEQ has clarified the language in the DSD where model fit is compared between the Cox proportional hazards model and the linear two-piece spline model (now in Appendix 6), to note that the AIC values are comparable between the two models. At the same time, it should be noted that TCEQ's Cox proportional hazards model is statistically shown to much more accurately predict lymphoid cancer in the primary NIOSH dataset, and is now also statistically shown to be accurate for a validation dataset (i.e., the UCC data), whereas the linear two-piece spline model is not.

Comment 2:

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.

Response to Expert 6 Comment 2:

The TCEQ appreciates Expert 6's comment about equitably-presented visual fits to categorical data and notes the expert's opinion that greater weight should be placed on formal goodness-of-fit tests. The TCEQ only discusses visual fit (and only in an Appendix) because of USEPA's reliance on it. By contrast, the TCEQ does not rely on visual model fit as a primary consideration for model choice, but rather principally relies on MOA and various statistical diagnostics of model fit (i.e., AIC and p-values, statistical analyses of model accuracy), consistent with the comment.

Comment 3:

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.

Response to Expert 6 Comment 3:

In the final DSD the discussion of visual fit has now moved to Appendix 6, Section A6.3.1.2. This section further explains the methods used by TCEQ to evaluate the background rates for the various models. Section A6.3.1.2.1.2 discusses the mathematics of this concept explicitly and notes that the USEPA also had a footnote on Figure 4-3 stating "the different models have different implicitly estimated baseline risks." (USEPA 2016). Lastly, contrary to one of the statements made in Expert 6's comments, the Cox model does not use an unexposed referent group. Briefly, as now discussed in Section 4.3 of the final DSD, the Cox proportional hazards model defines a risk set for every case (e.g., every cancer mortality from the specific cause), rather than needing a control (i.e., unexposed) group to derive the slope of the relative risk model. This is standard methodology for Cox proportional hazards modeling. The Cox modeling risk sets include all the individuals that are at risk at the time the case occurred (e.g., the time of the cancer mortality from the specific cause), both exposed and unexposed workers. The TCEQ's analysis uses the full risk set, including unexposed and exposed individuals, for every case in the NIOSH study, each possibly having more than 17,000 individuals in the risk set.

Comment 4:

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.

Response to Expert 6 Comment 4:

In the DSD, the TCEQ objects to a comparison of the dose-response model results to the categorical results. This is because while assessing model fit by visual inspection to the underlying *modeled* datapoints is a commonly used technique (e.g., USEPA 2012), the dose-response models being judged by visual fit to the categorical results were fit to *different* data, the more refined individual data. The USEPA should not have used the categorical modeling results (which are not the primary data) to visually evaluate the fit of models to other data (the

individual data) as though the cruder categorical data represent the true underlying dose-response (Valdez-Flores et al. 2013). The other concern with comparing to the dose-response shape of these categorical results is that the shape can change with different numbers of categories, as is shown in Valdez-Flores et al. (2013). Therefore, the TCEQ in this DSD chose not to rely on visual comparison to the quintile categorical modeling results.

In response to Expert 6's concerns that only the Cox regression was equalized to RR=1 at zero dose, the TCEQ has now presented figures in Section A6.3.1.2.1.1 showing the linear two-piece spline model (Figure 14) or the Cox regression (Figure 15) normalized to RR=1 at zero dose. As more fully discussed in the response above, contrary to the comment, the Cox model does not use an unexposed referent group.

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 2

Comment:

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.

Response to Expert 2 Comment:

The TCEQ appreciates Expert 2's comments on the lack of biological and mechanistic support for a supra-linear model, and notes that the expert asserts support for a linear model if the goal is a more conservative approach, and that there is justification for use of a less-than-linear approach based on the available biological data. These comments are consistent with the TCEQ's evaluation and approach.

Expert 3

Comment:

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.

Response to Expert 3:

The TCEQ appreciates Expert 3's comments about the burden of proof for using a supra-linear model, and notes that the expert thinks that there is not a convincing case, based on mechanistic or epidemiologic evidence, to support a supra-linear model. The TCEQ agrees with the Expert 3's comment that the shape of the dose-response curve will be based on inference from all available lines of evidence, including endogenous EtO levels and predicted population risk. The TCEQ DSD considers multiple lines of evidence that collectively and robustly support its dose-response model selection.

Expert 4

Comment 1:

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).

Response to Expert 4 Comment 1:

The TCEQ appreciates Expert 4's considerations that mechanistic data are needed to inform the low-dose extrapolation of the dose-response curve, and that, in the apparent absence of specific mechanistic data to support it, that a supra-linear model does not appear to be supported. These comments are buttressed by expert 3's comment that the burden of proof for invoking a supra-linear model is on those who postulate it.

Comment 2:

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?)

Response to Expert 4 Comment 2:

The TCEQ appreciates Expert 4's comments about the value of distinguishing between scientific and policy-based decisions. The final DSD includes an introduction (Section 2.1) that discusses the TCEQ guidance for decisions made based on the data, particularly based on MOA information. As is noted in Section 2.1 as well as in the MOA discussion in Section 3.2, as per TCEQ guidance (TCEQ 2015) either a mutagenic or an unknown MOA dictate a non-threshold approach to dose-response modeling (i.e., deriving a URF through linear low-dose extrapolation). As discussed in the context of the final DSD, under TCEQ guidance, the Cox proportional hazard model is preferred and using a supra-linear dose-response model must be justified. Robust MOA data are the primary data that could provide scientific justification for a supra-linear model under TCEQ guidelines (TCEQ 2015). While TCEQ can offer no examples of the adoption of an overall supra-linear model for epidemiological dose-response data, as discussed in the DSD, supra-linear responses are generally associated with an MOA that involves the saturation of metabolic activation where fewer electrophiles are formed per unit dose at higher exposures, which is not the case for EtO.

The TCEQ has substantially rewritten the DSD so that that main text is focused on the EtO data and the analyses, decisions, and justifications made by the TCEQ in deriving a unit risk factor for EtO. Because the TCEQ has conducted an independent analysis of the data, there is little discussion of a supra-linear dose-response model because there were no data in the MOA analysis that would suggest that a supra-linear model should be considered.

Expert 5

Comment 1:

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.

Response to Expert 5:

The TCEQ has edited the DSD to include the definition of "supra-linear" as a model with a dose-response curve that is steeper than linear as illustrated in Figure 2 of the final DSD where the low-dose slope is steep beginning at zero dose and then transitions at higher doses to a shallower slope (as is noted by Expert 5). By TCEQ's definition this can include curvilinear models or multi-part linear spline models with this same shape. The TCEQ now largely discusses the USEPA (2016) dose-response model as a linear two-piece spline, with a steeper slope in the low-dose region. As per the suggestions of the expert peer reviewers, much of the discussion of the USEPA's model has been removed from the main text of the DSD.

As noted in TCEQ's response to Expert 5's comments on charge question 3, because of the absence of literature and the nuanced considerations of gene mutations, the TCEQ does not find support for the hypothesis of gene mutations in carcinogenesis acting like a receptor occupation model (such as the Michaelis-Menton model). TCEQ has removed the Michaelis-Menton kinetics example from the DSD because while it is an example of a model with a mechanistic basis, like the Chemical Industry Institute of Toxicology formaldehyde model cited as an example, the TCEQ did not mean to suggest that this type of model may apply to a carcinogen with a direct acting mutagenic MOA (where no metabolic activation is required). This type of model is not an appropriate comparator for the dose-response relationship of a direct acting carcinogen to human cancer mortality.

The TCEQ has substantially rewritten the DSD so that the main text is focused on the EtO data and the analyses, decisions, and justifications made by the TCEQ in deriving a unit risk factor for EtO. Because the TCEQ has conducted an independent analysis of the data, there is no need to justify a change from the EPA model. There is still some discussion of the different choices made by EPA compared to TCEQ, but these are largely restricted to the Response to Public Comments document and Appendix 6 of the DSD.

The TCEQ DSD lays out important scientific considerations and analyses that are convincing and favor the selection of the Cox proportional hazards model over the linear two-piece spline model. These include: 1) better alignment with the MOA for a direct-acting mutagenic carcinogen; 2) parsimony (i.e., fewer model components); and 3) accuracy in predicting lymphoid cancer in the primary NIOSH cohort dataset as well as in a validation analysis using the UCC cohort dataset. These facts show that the models are not equal as implied, but rather favor the use of the Cox proportional hazards model, which is why the TCEQ chose this model. The TCEQ agrees with expert 3's comment that the burden of proof for invoking a supra-linear model is on those who postulate it, which is consistent with TCEQ guidelines (TCEQ 2015).

Expert 6

Comment 1:

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., ~$dose^2$). 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, $P(d) = 1- \exp(-a-bd)$, 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.

Response to Expert 6 Comment 1:

The TCEQ has edited the DSD to include the definition of "supra-linear" as a model with a dose-response curve that is steeper than linear as illustrated in Figure 2 of the final DSD where the low-dose slope is steep beginning at zero dose and then transitions at higher doses to a shallower slope (as is noted by Expert 5). By TCEQ's definition this can include curvilinear models or multi-part linear spline models with this same shape. The TCEQ now largely discusses the USEPA (2016) dose-response model as a linear two-piece spline, with a steeper slope in the low-dose region. Much of the discussion of the USEPA's model has been removed from the main text of the DSD.

Comment 2:

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.

Response to Expert 6 Comment 2:

The TCEQ acknowledges that both the shape and potency of the dose-response curve are important factors in the dose-response analysis. As noted above, the TCEQ has now defined

"supra-linear" as a curve that is steeper at low doses than at higher doses, and we now largely discuss the USEPA (2016) dose-response model as a linear two-piece spline, with a steeper slope in the low-dose region. The potency aspect of the analysis is addressed by the reality-checks discussed elsewhere in this document, which support the accuracy of the Cox proportional hazards model over the linear two-piece spline model (e.g., see Section 4.2.3 of the DSD).

Comment 3:

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.

Response to Expert 6 Comment 3:

The TCEQ has edited the DSD to include the definition of "supra-linear" as a model with a dose-response curve that is steeper than linear as illustrated in Figure 2 of the final DSD where the low-dose slope is steep beginning at zero dose and then transitions at higher doses to a shallower slope (as is noted by Expert 5). By TCEQ's definition this can include curvilinear models or multi-part linear spline models with this same shape. The TCEQ now largely discusses the USEPA (2016) dose-response model as a linear two-piece spline, with a steeper slope in the low-dose region. Much of the discussion of the USEPA's model has been removed from the main text of the DSD, as has the quote from Crump and Allen (1985).

Comment 4:

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 exogeneous exposures (i.e., at zero exogeneous exposure) will be positive. Therefore, the dose-response as a function of exogeneous 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., ≤ 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 exogeneous exposures. EPA did not claim that the response to exogeneous doses should be sub-linear, in fact, as suggested in the previous paragraph, they argued that the response to exogeneous should be linear (linear, as defined in Crump 2011).

Response to Expert 6 Comment 4:

As noted above, the TCEQ stated that "EPA considers it highly plausible that the dose-response relationship over the endogenous range is sublinear." While Expert 6 protests this statement because EPA is discussing the relationship in the endogenous range, the TCEQ notes that the quoted text from USEPA does state that the sublinearity occurs over the endogenous range. While it is true that by definition exogenous exposure will be in addition to endogenous exposure, there is no *a priori* reason to expect that the sublinear dose-response has an upward curve that begins precisely at the point of additional exogenous exposure (as the curve does in Figure 1 of Crump 2014). Moreover, as the general population has a distribution of endogenous levels, this variability makes identification of the interface between endogenous and exogeneous exposures nearly impossible. It is reasonable to assume that the DNA repair processes that operate for endogenous damage will still be able to operate with a certain amount of exogenous damage, particularly if those levels of exogenous damage are in the

background air concentration range. For example, exogenous environmental doses of EtO are commonly low enough that total internal dose (due to endogenous + exogenous) would still be within the normal endogenous range (where USEPA indicates sublinearity applies) for a large portion of the population (e.g., urban background concentrations of ≤0.2 ppb are below even the 1st percentile of the estimated endogenous distribution that corresponds to a continuous air concentration of ≈0.37 ppb, and are substantially below the endogenous mean corresponding to ≈1.9 ppb in air; Kirman and Hays 2017) . In addition, it seems that just because a slope is positive, does not make the shape of the curve linear (as evidenced by Figure 1 of Crump 2017, which is sublinear in shape).

Comment 5:

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 exogeneous exposures are added to the endogenous exposures. Therefore, I see no basis for a direct comparison since the exogeneous exposures are added to endogenous exposures and never occur alone.

Response to Expert 6 Comment 5:

The referenced figure was extraneous and unnecessary and has been removed from the final DSD, which is much more focused on the critical scientific components of the TCEQ's hazard and dose-response assessment.

Charge Question 8

As summarized in DSD [draft dated 1/31/20] 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 2

Comment:

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.

Response to Expert 2:

The TCEQ notes Expert 2's comment that the model predictiveness and reality checks are reasonably done, all of which support the model selected by the TCEQ. The TCEQ has substantially rewritten the DSD so that that main text is focused on the EtO data and the

analyses, decisions, and justifications made by the TCEQ in deriving a unit risk factor for EtO. Most of the comparison to the USEPA model has been removed, as well as much of the language criticizing EPA.

Expert 3

Comment:

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.

Response to Expert 3:

The TCEQ notes Expert 3's comments that the model predictiveness and reality checks are well-justified and clearly presented. The TCEQ agrees with the critical importance of seeking multiple approaches to judge whether or not a dose-response model and subsequent URF are sensible, which is why the DSD utilizes multiple lines of evidence to robustly evaluate its modeling approach (e.g., MOA, model accuracy analyses for the NIOSH cohort as well as a validation analysis of the NIOSH-based model using UCC data).

Expert 4

Comment 1:

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).

Response to Expert 4 Comment 1:

While TCEQ agrees that it would be optimal to ground-truth the risk estimates based on dose-response relationships across the toxicological and epidemiology database, it is not clear that at this time the data or methods are available for this additional analysis. However, in regard to ground-truthing using additional epidemiological data, the final DSD contains a new validation analysis where the UCC data were used to show that TCEQ's model based on the NIOSH cohort is also predictive of lymphoid cancer deaths in the UCC cohort, whereas the linear two-piece

spline model is statistically significantly over-predictive (also see response to Expert 4 Comment 2).

Comment 2:

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?

Response to Expert 4 Comment 2:

The TCEQ appreciates this very helpful suggestion by Expert 4, and the agency undertook this analysis with other epidemiology study cohorts. In a validation analysis, the TCEQ used the UCC cohort data to evaluate the predictiveness of the Cox proportional hazards model and the linear two-piece spline model, both having been based on the NIOSH data. Despite substantial differences in the exposure assessments for the NIOSH and UCC cohorts (see Section 4.1.2 of this DSD and Section 4.1 of USEPA 2016), using UCC cohort data to evaluate the validity of the models derived based on the NIOSH dose-response assessment results in the same conclusion; namely that the Cox proportional hazards model is reasonably accurate at predicting the number of lymphoid cancer mortalities observed in the UCC cohort while the linear two-piece spline model is statistically significantly over-predictive whether using the maximum likelihood estimate (MLE) or upper bound (see Section A3.3.3 of Appendix 3). Thus, the Cox model is demonstrated to be reasonably predictive and realistic, lending strong support to its scientific credibility for regulatory agency use (e.g., EtO URF derivation).

Comment 3:

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 haematological 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

Response to Expert 4 Comment 3:

The TCEQ has reviewed the Environment Canada (2001) EtO assessment and noted the comparison of risks between the animal and human study results. However, as noted by that assessment, the limitations of these comparisons do not allow quantitative comparisons of risks, and therefore we have not conducted such analyses in this DSD.

Expert 5

Comment 1:

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.

Response to Expert 5 Comment 1:

The TCEQ appreciates Expert 5's identification of an important point: the question of whether low-dose extrapolation should occur based on observations from much higher doses (thousands or millions of times higher, in this case). Both agencies have done exactly that, used a high dose dataset to estimate the low-dose linear part of the dose-response curve, so TCEQ does not seek to rebut the standard regulatory agency practice of linear low-dose extrapolation based on high dose data (whether epidemiological or from animal studies). The only difference is the model used to do it. The TCEQ has concluded, based on the likely mutagenic MOA of EtO, that there should be a low-dose linear extrapolation from high dose data, using the Cox proportional hazards model as a standard model for that purpose. This is supported by multiple lines of evidence including statistical analyses showing that TCEQ's model based on the NIOSH cohort accurately predicts lymphoid cancer deaths both in the NIOSH cohort and a model validation dataset (the UCC cohort), while the corresponding linear two-piece spline model is statistically significantly over-predictive for both datasets. Moreover, the TCEQ agrees with Expert 3's comment that the burden of proof for invoking a supralinear model is on those who postulate it, which is consistent with TCEQ guidelines (TCEQ 2015).

The TCEQ has substantially rewritten the DSD so that the main text is focused on the EtO data and the analyses, decisions, and justifications made by the TCEQ in deriving a unit risk factor for EtO. Because the TCEQ has conducted an independent analysis of the data, there is no need to justify a change from the USEPA model. There is still some discussion of the different choices made by USEPA compared to TCEQ, but these are largely restricted to the Response to Public Comments document and Appendix 6 of the DSD.

The TCEQ has included in the DSD a more thorough discussion of the potential early key events in EtO-carcinogenesis in humans, and notes in Chapter 4 and Appendix 1 that while this information is useful for the MOA determination, it is not robust enough to derive a chronic carcinogenicity toxicity factor.

Comment 2:

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?

Response to Expert 5 Comment 2:

As noted above, the TCEQ has substantially rewritten the DSD so that that main text is focused on the EtO data and the analyses, decisions, and justifications made by the TCEQ in deriving a unit risk factor for EtO. Section 4.2 of the final DSD addresses the parameters that the TCEQ used to consider dose-response model choice, including MOA information, AIC and p-value, and the model accuracy when compared to the NIOSH cohort data that is the data source for the dose-response model. As to accuracy, for example, the models do not perform the same. The final DSD includes statistical analyses showing that the Cox proportional hazard model based on the NIOSH cohort accurately predicts lymphoid cancer deaths both in the NIOSH cohort and a model validation dataset (the UCC cohort), while the linear two-piece spline model is statistically significantly over-predictive for both datasets.

Comment 3:

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."

Response to Expert 5 Comment 3:

The TCEQ has modified the language cited above about the fairness of the evaluation, but we do note here that a model should at minimum be able to predict the data on which it is based. In regard to the comment that a fairer evaluation would be to compare the predictions to data that were not used to estimate the parameters of the model, the TCEQ has now done exactly that. The final DSD contains a new validation analysis where the UCC data were used to evaluate the dose-response models that were based on the NIOSH cohort, and this analysis showed that the Cox proportional hazards model is also predictive of lymphoid cancer deaths in the UCC cohort, whereas the linear two-piece spline model is statistically significantly over-predictive.

Comment 4:

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.

Response to Expert 5 Comment 4:

The methods for predicting the number of cases from the NIOSH cohort inputs is provided in Appendix 3 Sections 3.4 and 3.5. These sections now provide a far more robust explanation of the TCEQ's reality check calculations and provide hypothetical examples to walk readers through the mathematics. Aside from the mathematics, Figure 2 below (which illustrates the Cox proportional hazard and linear two-piece spline model fit to the individual NIOSH cohort lymphoid mortality data) shows how steep the linear two-piece spline model is at low doses, which generates the over-estimates of predicted lymphoid cancer deaths.

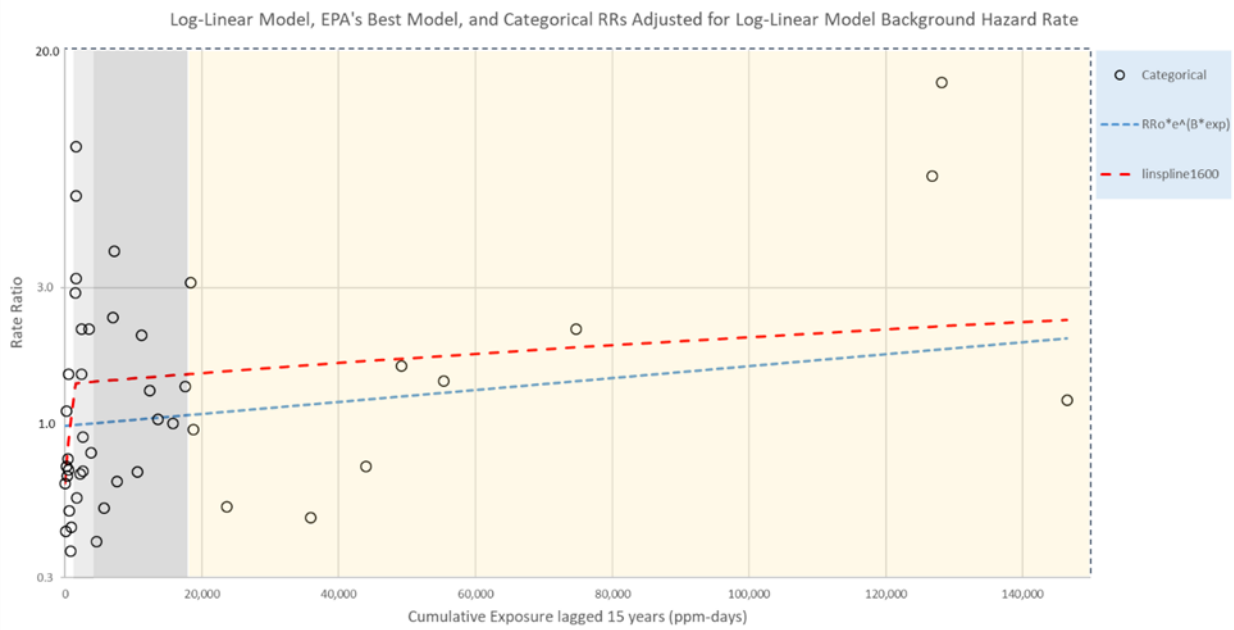

Figure 2. Lymphoid cancer death categorical risk ratios and the Cox proportional hazards and linear two-piece spline ("knot" at 1,600 ppm × days) fitted models (Figure 15 in the final DSD).

Comment 5:

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.

Response to Expert 5 Comment 5:

The methods for determining the confidence bounds on the predicted numbers for the model prediction are explained in Appendix 3 Section 3.3.

A confidence interval on the standardized mortality ratio (SMR) can be constructed as per Rothman and Boice (1979), who use the following equations to derive 100(1-α)% confidence intervals for the SMR:

$$SMR_{LCL} = \frac{Observed}{Expected} \times \left(1 - \frac{1}{9 \times Observed} - \frac{Z_{\alpha/2}}{3 \times \sqrt{Observed}}\right)^3$$

and

$$SMR_{UCL} = \frac{(Observed + 1)}{Expected} \times \left(1 - \frac{1}{9 \times (Observed + 1)} + \frac{Z_{\alpha/2}}{3 \times \sqrt{Observed + 1}}\right)^3$$

where SMR_{LCL} is the 100(1-α/2)% lower confidence limit on the SMR, SMR_{UCL} is the 100(1-α/2)% upper confidence limit on the SMR, *Observed* is the number of observed cause-specific deaths (e.g., lymphoid cancer deaths) in the study $(i.e., Observed = \sum_i y_{oi})$, *Expected* is the expected cause-specific deaths (e.g., lymphoid cancer deaths) derived from the reference population background rates $\left(i.e., Expected = \sum_i p_{oi} \frac{y_{ri}}{p_{ri}}\right)$, and $Z_{(\alpha/2)}$ is the 100(1- α/2)% percentile of the standard normal distribution. The 100(1-α)% confidence interval for an SMR is given by the interval (SMR_{LCL}, SMR_{UCL}).

Expert 6

Comment 1:

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).

Response to Expert 6 Comment 1:

The TCEQ acknowledges that both the Cox model and the EPA model are linear in the low-dose region, and we have now included a definition of "supra-linear" in the DSD.

Comment 2:

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.

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.

Response to Expert 6 Comment 2:

TCEQ agrees that the best data for predicting low-dose response to low-dose EtO exposure are low-dose data. However, all the occupational and animal study data for EtO are at least thousands of times higher than low-dose (i.e., environmental) exposures. Therefore, TCEQ does not in this case define "lowest available dose" as equivalent to "low-dose" and cannot justify relying exclusively or differentially on the lowest dose occupational data when choosing the shape of the dose-response curve. In the absence of other information, the TCEQ chose to use the putative mutagenic MOA to inform a low-dose linear extrapolation, and we used the standard Cox proportional hazards model for the EtO dose response. One method that TCEQ used to test the models was to assess how accurately the TCEQ and the USEPA models predicted the NIOSH cohort data that were used to derive the model. In so doing we found that the TCEQ model can predict the number of lymphoid cancer deaths with 95% confidence overall (Tables 6 and 29) and also in every exposure quintile including the lowest quintile of exposures of the NIOSH study (Tables 7 and 30). In contrast, other proposed models over-predict (at the 5% significance level) the overall number of lymphoid cancer deaths in the NIOSH study (Tables 6 and 29) and also the number of lymphoid deaths in the lowest quintile of EtO exposures and other quintiles (Tables 7 and 30). This suggests that the TCEQ's model provides the best fit to all the data, including the lowest dose data.

Comment 3:

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.

Response to Expert 6 Comment 3:

The DSD now further discusses the choice of estimates of EtO from Kirman and Hays (2017) in Appendix 5. Appendix 5 discusses that the EtO estimates chosen by TCEQ from Kirman and Hays (2017) appear reasonable considering: (1) the geometric mean HEV levels reported for nonsmokers (≈31 pmol/g) and smokers (≈143 pmol/g) by Jain (2020) based on 2013-2016 NHANES data (see Table 3 of Jain 2020); and (2) the background HEV levels in control rats (≈42-50 pmol/g Hb) and mice (≈58-100 pmol/g Hb) (Walker et al. 1993, 2000). The estimates referred to are in Appendix 6 of the final DSD and are not a consideration in model selection.

Comment 4:

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).

Response to Expert 6 Comment 4:

In the final DSD, the term "biologically meaningful" is used only once, in an Appendix 5 discussion of biological context. While even an infinitesimal increase in a chemical in a biological system could be considered to produce a very small increase in risk because of the choice of a linear risk model, there is still value and importance in discussing how that amount of chemical would be handled in a biological system. However, the concept of biologically meaningful or significant was not relied upon in the URF development process documented in the final DSD.

Comment 5:

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 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.

Response to Expert 6 Comment 5:

The TCEQ appreciates Expert 6's clarification of what is meant by "low dose data" and the potential method provided to identify this data in the occupational cohort.

For epidemiological studies with individual exposure history (such is the case for the NIOSH study) the data are not specified by dose group as is the case in animal studies. In fact, in epidemiological study analyses with individual exposure data the experimental "unit" is not the individual person, but rather the person time (e.g., person year). That is, instead of thinking that in the NIOSH study there are about 17,000+ workers it is more appropriate that there are about 500,000 person years. (and yet, that is a simplified approximation of how epidemiological data are treated). Thus, defining "low dose" would mean including the person-time with cumulative exposures less than or equal to a cut-off (and that could be done). What is difficult to ascertain is how to determine if the data conform to a "linear" exposure-response relationship. Even for the full data set, it is difficult to ascertain whether the model is linear or not.

In addition, the method provided seems to not necessarily identify low dose data, but rather identifies that part of the data at the low dose end that fits a single linear response curve in the occupationally-exposed cohort. However, even if this analysis is conducted it will still not be clear if the identified linear model is more appropriate to fit the orders-of-magnitude lower ambient exposures of concern than a model that is derived based on all of the occupational data. Therefore, the TCEQ has decided to continue to base the dose-response model on the entire cohort of data.

Charge Question 9

In DSD [draft dated 1/31/20] 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 [draft dated 1/31/20] 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 2

Comment 1:

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.

Response to Expert 2 Comment 1:

The TCEQ appreciates Expert 2's evaluation and notes their agreement that the modeling choices and general descriptions of rationales are acceptable.

Comment 2:

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 applied when assessing the risks to the general population (not workers) exposed to EtO, but this is not made clear in the DSD.

Response to Expert 2 Comment 2:

The TCEQ has now noted in the DSD that the URF is ADAF-adjusted because this URF is being applied to the general population.

Comment 3:

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.

Response to Expert 2 Comment 3:

The TCEQ has substantially rewritten the DSD so that that main text is focused on the EtO data and the analyses, decisions, and justifications made by the TCEQ in deriving a unit risk factor for EtO.

Comment 4:

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.

Response to Expert 2 Comment 4:

The TCEQ is principally considering the breast cancer data in a WOE evaluation for the hazard assessment question of whether EtO causes breast cancer. Although the TCEQ has ultimately chosen to designate EtO as a likely human carcinogen, this does not mean that all human cancers are caused by EtO. The primary types of cancer with some evidence in human studies are lymphoid cancers and breast cancer. As part of the process for deciding if there was likely to be a causal relationship between EtO and breast cancer in humans, the TCEQ investigated the various available data, which included animal studies showing possible increases in mammary carcinogenesis with EtO exposure. However, the IARC (2019) work demonstrates that the state of the science does not support tumor site concordance as a general principle. Thus, the positive mammary carcinogenesis results in one animal species (mice but not rats) do not provide support for a specific site of cancer in humans, and therefore does not inform the breast cancer decision outside of providing information about the initial carcinogenesis determination. This information has now been clarified in the DSD in Section 3.3.1.1.1.

Comment 5:

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.

Response to Expert 2 Comment 5:

The TCEQ shares Expert 2's concerns about the evidence for carcinogenicity of EtO to humans and has chosen to adopt the designation of *likely to be carcinogenic to humans.* Section 3.3 of the final DSD now provides a more thorough discussion and WOE of the hazard determination for EtO and human cancer.

Comment 6:

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 × 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. [Note in Text: "What does this mean" attached to the text "qualitatively similar results"].

Response to Expert 2 Comment 6:

The 1 in 100,000 excess risk level is TCEQ policy and is documented in the TCEQ guidelines (TCEQ 2015), as are weighting procedures to use when more than one epidemiological study is used for URF derivation. "Qualitatively similar results" was essentially meant to convey that based on either weighting factor the conclusion drawn was going to be similar, namely that the URF was overwhelmingly going to be based on the NIOSH cohort results. This language has been removed from the DSD.

Comment 7:

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.

Response to Expert 2 Comment 7:

The TCEQ has replaced the word "conservative" with "health-protective" in this section.

Comment 8:

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?

Response to Expert 2 Comment 8:

The statement about the numbers being within a factor of 1.3 has been removed.

Comment 9:

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.

Response to Expert 2 Comment 9:

The TCEQ has edited this statement to: *"USEPA (2016) indicates that there are no data on the relative susceptibility of children (or young animals of other species) to EtO…"*

Comment 10:

3.4.2 Evaluating Susceptibility from Early-Life Exposures

*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 µg/m*** ^{3,} ***[emphasis in original text]** (i.e., 2.5E-06 per ppb × 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*.

Response to Expert 2 Comment 10:

The discussion of the ADAF application to the URF has been expanded and is found in Section 4.3.3.1 of the final DSD. The discussion of the effects of an additional 10-fold ADAF for *in utero* exposure has been removed from the DSD as it simply refers to a peripherally-related topic not part of TCEQ guidance (TCEQ 2015).

Comment 11:

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.

Response to Expert 2 Comment 11:

The Kirkeleit et al. (2013) paper is the citation for the lack of healthy worker effect for lymphoid cancer in general. The findings from this paper are discussed in the final DSD in Section 3.1.1.2. More specific to the assessment, the DSD also documents the lack of such an effect based on results of the NIIOSH study (see Section A3.3.1 of the final DSD).

Comment 12:

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?

Response to Expert 2 Comment 12:

This is the assumption made in the models fit to the NIOSH data. The rate ratios are the ratio of the hazard rate at a given exposure to the background hazard rate. This information has been added to the referenced text in the DSD.

Comment 13:

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.

Response to Expert 2 Comment 13:

This section of the DSD (Appendix 2 in the revised DSD, Appendix 3 in the final DSD) describes the model evaluation that the TCEQ uses to determine if the Cox proportional hazards model can successfully predict the dataset that the model is based on. This model evaluation does not directly affect the URF. The conclusions to draw from this appendix are that the TCEQ model can reasonably accurately predict the responses that were observed as opposed to other models that cannot reasonably accurately reproduce the observed data (i.e., they statistically significantly over-predict).

Comment 14:

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?

Response to Expert 2 Comment 14:

While this comparison was originally made for biological context, it has been removed from the final DSD.

Comment 15:

> *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.

Response to Expert 2 Comment 15:

The word "recalculated" is now used in the DSD in place of the word "corrected" in referenced to the AIC and p-values.

Comment 16:

> *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.

Response to Expert 2 Comment 16:

The TCEQ now explicitly notes that the conclusions are those of TCEQ based on the analyses presented in the DSD.

Comment 17:

> *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.

Response to Expert 2 Comment 17:

In the DSD, the TCEQ objects to a comparison of the dose-response model results to the categorical results. This is because while assessing model fit by visual inspection to the underlying *modeled* datapoints is a commonly used technique (e.g., USEPA 2012), the dose-response models being judged by visual fit to the categorical results were fit to *different* data, the more refined individual data. The USEPA should not have used the categorical modeling results (which are not the primary data) to visually evaluate the fit of models to other data (the individual data) as though the cruder categorical data represent the true underlying dose-response (Valdez-Flores et al. 2013). The other concern with comparing to the dose-response shape of these categorical results is that the shape can change with different numbers of categories, as is shown in Valdez-Flores et al. (2013). Therefore, the TCEQ in this DSD chose not to rely on visual comparison to the quintile categorical modeling results.

Expert 3

Comment:

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.

Response to Expert 3:

The TCEQ appreciates Expert 3's evaluation and notes their agreement that TCEQ's modeling assumptions and calculations are fully documented and supported and that the choices seem reasonable. In addition, there are several sensitivity analyses in the DSD including the consideration of the impact of a healthy worker effect on the dose-response model predictions (Sections 4.2.3 and A3.3.2), and testing the model predictions on the UCC cohort using the dose-response model built from the NIOSH cohort (Sections 4.2.3 and A3.3.3).

Expert 4

Comment 1:

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 in relatively consistent with a mutagenic mode of action).

Response to Expert 4 Comment 1:

TCEQ appreciates Expert 4's analysis and notes their agreement that support for causality between EtO and breast cancer is limited. Further discussion of Expert 4's concerns can be found in response to charge question 4.

Comment 2:

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.

Response to Expert 4 Comment 2:

The TCEQ determined in the hazard assessment that there is inadequate evidence that EtO causes breast cancer, and therefore did not further assess whether the lymphoid cancer assessment was protective of breast cancer.

Expert 5

Comment 1:

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.

Response to Expert 5 Comment 1:

The TCEQ appreciates Expert 5's evaluation and notes the experts lack of disagreement with the URF and ADAF calculations. As noted above, the TCEQ has substantially rewritten the DSD so that the main text is focused on the EtO data and the analyses, decisions, and justifications made by the TCEQ in deriving a unit risk factor for EtO. Because the TCEQ has conducted an independent analysis of the data, there is no need to justify a change from the USEPA model. There is still some discussion of the different choices made by USEPA compared to TCEQ, but these are largely restricted to the Response to Public Comments document and Appendix 6 of the DSD.

Regulatory agencies may err on the side of conservatism when other scientific considerations are equal, but in this case the scientific considerations are not equal. The TCEQ has used

multiple lines of evidence as a basis for selection of the Cox proportional hazards dose-response model rather than the linear two-piece spline model. These include: 1) better alignment with the MOA for a direct-acting mutagenic carcinogen; 2) parsimony (i.e., fewer model components); and 3) accuracy in predicting lymphoid cancer in the primary NIOSH cohort dataset as well as in a validation analysis using the UCC cohort dataset. These facts show that the models are not equal as implied, but rather favor the use of the Cox proportional hazards model, which is why the TCEQ chose this model.

Comment 2:

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.

Response to Expert 5 Comment 2:

As noted in response to Expert 5's comments on charge question 3, the TCEQ will continue to investigate the suggested methods for future use but at this time does not consider that these methods can be applied to this EtO carcinogenicity dose-response assessment. See the previous response and the final DSD regarding results of the evaluation of model accuracy.

Expert 6

Comment:

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.

Response to Expert 6:

The TCEQ has edited the DSD to more clearly explain the application of ADAFs in the URF calculation in Section 4.3.3.1.

Charge Question 10

Based on biomarker data, various sections of the DSD [draft dated 1/31/20] (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 2

Comment:

I generally found the material to be appropriate. See some notes on marked-up copy of the DSD.

> *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.

Response to Expert 2:

The TCEQ appreciates Expert 2s evaluation and notes the expert's opinion that the endogenous EtO analyses are appropriate.

The referenced statement is based on the calculations of the NIOSH cohort. These calculations suggest that if USEPA's linear two-piece spline model is correct (assuming both a steep lower-dose region followed by a less-steep higher dose region), then there should be an observation of increased cancers in the lower occupational exposure groups. Results of the model accuracy analysis for lymphoid cancer mortality in Table 30 of the final DSD provide the statistical basis for this statement.

Expert 3

Comment:

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.

Response to Expert 3:

The TCEQ appreciates Expert 3’s evaluation and notes the expert’s opinion that the TCEQ’s analyses are quite helpful and seem to be scientifically appropriate. While the analyses do raise serious concerns about the use of the linear two-piece spline model, it should be noted that the TCEQ’s analyses of endogenous and background EtO levels were not used in model selection. For that, the TCEQ relied on MOA and other considerations (e.g., model fit criteria, evaluations of the accuracy of model predictions, described in Section 4.2 of the final DSD).

Expert 4

Comment 1:

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.

Response to Expert 4 Comment 1:

The discussion of endogenous EtO in the final EtO DSD is restricted to Appendix 5 and Appendix 6.4.2, where it provides some biological context for understanding the comparative levels of EtO from endogenous and exogenous sources. This is consistent with guidance from USEPA (2005a) that indicates that biomarkers of internal dose (such as the EtO hemoglobin N-(2-hydroxyethyl)-valine (HEV) adducts) can provide insight into the potential shape of the dose-response curve at doses below those at which tumors are induced experimentally. Appendix 5 compares the HEV adduct level changes that would occur with a continuous exposure to 4 ppb EtO (the 1 in 100,000 excess risk concentration based on the ADAF-unadjusted URF) to endogenous levels of EtO, and finds that an additional ≈43.6 pmol/g Hb due to continuous exogenous exposure to 4.0 ppb would be predicted to:

- Increase the HEV level for the median non-smoker to between the 95th and 99th percentiles of normal endogenous background levels; and
- Increase the HEV level in 90th percentile non-smokers to over the 99th percentile.

This information provides biological context for the TCEQ's 1 in 100,000 excess risk concentration, but it is not a consideration for model choice. The text no longer includes the quote from Swenberg et al. (2008).

Comment 2:

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?

Response to Expert 4 Comment 2:

The discussion of endogenous conversion of exogenous ethylene to EtO is no longer included in the final DSD. As noted above, the TCEQ now restricts the discussion of endogenous EtO to Appendix 5, where it provides biological context for the TCEQ's 1 in 100,000 excess risk concentration, and Appendix A6.4.2. In Section A6.4.2 the TCEQ applies the USEPA (2016) URF, which was developed based solely on occupational exposure where exposure to the chemical in ambient air outside the workplace was part of background exposure, to endogenous and background EtO levels based on the toxicological principle that equal doses give rise to equal risk. That principle of considering equal internal doses as equipotent in producing carcinogenic effects allowed for a rough estimation of the lymphoid tumor risk in the population given endogenous and background doses, using USEPA (2016)'s URF. This analysis was not used in model selection or to derive the TCEQ URF, but only to provide a reality check when assessing the USEPA's URF.

Comment 3:

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.

Response to Expert 4 Comment 3:

In response to this and other expert's suggestions, the TCEQ has substantially rewritten the DSD so that the main text is focused on the EtO data and the analyses, decisions, and justifications made by the TCEQ in deriving a unit risk factor for EtO. Sections of peripherally-related materials deemed extraneous were deleted and do not appear in the final DSD. Because the TCEQ has conducted an independent analysis of the data, there is no need to justify a change from the USEPA model. There is still some discussion of the different choices made by USEPA compared to TCEQ, but these are largely restricted to the Response to Public Comments document and Appendix 6 of the DSD. The final DSD also now contains an introductory chapter (Chapter 2) that explicitly states the problem formulation for the DSD.

Comment 4:

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).

Response to Expert 4 Comment 4:

The TCEQ appreciates Expert 4's suggestions about methods for ground-truthing TCEQ's approach and has implemented some of those suggestions as discussed elsewhere in this document.

Expert 5

Comment:

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.

Response to Expert 5 Comment:

The TCEQ's dose-response assessment does not rely on the consideration of endogenous EtO levels. Rather, the TCEQ relied on MOA and other considerations such as model fit criteria combined with evaluations of the accuracy of model predictions for both the NIOSH and UCC cohorts. It is the weight of evidence from these multiple considerations that individually and together robustly support TCEQ's model. The discussion of endogenous EtO in the final EtO DSD is restricted to Appendix 5 and Appendix 6.4.2, where it provides some biological context for understanding the comparative levels of EtO from endogenous and exogenous sources. This is consistent with guidance from USEPA (2005a) that indicates that biomarkers of internal dose (such as the EtO hemoglobin N-(2-hydroxyethyl)-valine (HEV) adducts) can provide insight into the potential shape of the dose-response curve at doses below those at which tumors are induced experimentally. Appendix 5 compares the HEV adduct level changes that would occur with a continuous exposure to 4 ppb EtO (the 1 in 100,000 excess risk concentration based on the ADAF-unadjusted URF) to endogenous levels of EtO, and finds that an additional ≈43.6 pmol/g Hb due to continuous exogenous exposure to 4.0 ppb would be predicted to:

- Increase the HEV level for the median non-smoker to between the 95th and 99th percentiles of normal endogenous background levels; and

- Increase the HEV level in 90th percentile non-smokers to over the 99th percentile.

This provides context for the TCEQ's 100,000 excess risk concentration.

As noted in response to Expert 4's comments, the TCEQ in section A6.4.2 also does a rough estimation of the predicted lymphoid cancer rate in the general population based on the USEPA (2016) EtO URF. Those estimates generate predictions of > 3% lymphoid cancer from EtO exposure alone, which is considerably higher than the excess lifetime probabilities of 1 in 10,000 noted by Expert 5 and higher than the background rate cited by USEPA (2016). While these estimates do not impact TCEQ's model choice or the derivation of the TCEQ EtO URF, they do provide a reality check for derived EtO URFs.

Expert 6

Comment 1:

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 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."

Response to Expert 6 Comment 1:

In the final DSD, the term "biologically meaningful" is used only once, in an Appendix 5 discussion of biological context. The TCEQ thinks that both risk and biological considerations should be used when judging the suitability of a hazard and dose-response assessment. While even an infinitesimal increase in a chemical in a biological system could be considered to have a very small increase in risk because of the choice of a linear risk model, there is still value and importance in discussing how that amount of chemical would be handled in a biological system. However, the concept of biologically meaningful or significant levels of EtO was not relied upon in the URF development process documented in the final DSD.

Other aspects of this comment are addressed in responses to Experts 4 and 5, above.

Comment 2:

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.

Response to Expert 6 Comment 2:

The DSD now further discusses the choice of estimates of EtO from Kirman and Hays (2017), using supporting data from a new study of the National Health and Nutrition Examination Survey by Jain (2020).

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

Comment:

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

Response to Expert 1:

The TCEQ appreciates Expert 1's analysis of the epidemiologic findings and notes the expert's comment that these findings are presented clearly and objectively.

Expert 2

Comment 1:

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.

Response to Expert 2 Comment 1:

The TCEQ has substantially revised the DSD to improve clarity, logic, and transparency of the information collected, the analyses, and the conclusions.

Comment 2:

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.

Response to Expert 2 Comment 2:

The TCEQ has added definitions and/or plain language descriptions of the mathematics in those sections of the DSD indicated by Expert 2.

Comment 3:

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).

Response to Expert 2 Comment 3:

The TCEQ has added references and/or attributions in those sections of the DSD indicated by Expert 2.

Comment 4:

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.

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).

Response to Expert 2 Comment 4:

The TCEQ has substantially reorganized and revised the DSD (including the MOA section and discussion of cancer classification) to improve clarity, logic, and transparency of the information collected, the analyses, and the conclusions. The TCEQ recognizes the use of animal data for the cancer hazard assessment. At the same time, the TCEQ tries to convey that since current science does not support interspecies site concordance (IARC 2019), animal results are of limited use in supporting a particular potential site of EtO-induced carcinogenesis in humans.

Comment 5:

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.

Response to Expert 2 Comment 5:

The TCEQ has substantially revised the DSD to remove unnecessary use of emphasis in the text. Most of the comparison to the USEPA model has been removed, as well as much of the language criticizing USEPA. The result is a final DSD that focuses on the critical, multiple lines of evidence that individually and cumulatively robustly support TCEQ’s assessment.

Comment 6:

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.

Response to Expert 2 Comment 6:

The TCEQ has substantially revised the DSD, including Chapter 2, to improve organization and logical flow. The information about cancer incidence that was included in Chapter 2 has been removed in the final DSD.

Comment 7:

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.

Response to Expert 2 Comment 7:

In the final DSD the TCEQ has ensured that acronyms/abbreviations are defined at first use, and that the term “the agency” is replaced with a more specific label if the referenced agency is not clear in the text.

Expert 3

Comment 1:

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.

Response to Expert 3 Comment 1:

The TCEQ has substantially revised the DSD to improve clarity, logic, and transparency of the information collected, the analyses, and the conclusions. This revision better develops arguments fully in single locations in the document. Concerns with the USEPA approach are largely de-emphasized in the main body and now appear in the last appendix of the final DSD.

Comment 2:

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.

Response to Expert 3 Comment 2:

The TCEQ has substantially revised the DSD to ensure, as Expert 3 suggests, a tone of calm reason and careful consideration of the evidence. As mentioned above, concerns with the USEPA approach are largely de-emphasized in the main body and now appear in the last appendix of the final DSD.

Expert 4

Comment 1:

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.

Response to Expert 4 Comment 1:

The TCEQ has substantially rewritten the DSD so that the main text is focused on the EtO data and the analyses, decisions, and justifications made by the TCEQ in deriving a unit risk factor for EtO. Most of the comparison to the USEPA model has been removed, as well as much of the language criticizing USEPA. Concerns with the USEPA approach largely only appear in the last appendix of the final DSD.

Comment 2:

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 ????.

Response to Expert 4 Comment 2:

The TCEQ has substantially revised the DSD, including Chapter 2, to improve organization and logical flow. The references to risk communication and societal consequences as well as the information about cancer incidence that was included in Chapter 2 have been removed from the DSD.

Comment 3:

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.

Response to Expert 4 Comment 3:

The final DSD now contains an introductory chapter (Chapter 2) that explicitly states the problem formulation for the DSD. Discussion of the USEPA (2016) assessment mainly appears in the last appendix of the final DSD.

Expert 5

Comment:

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.

Response to Expert 5 Comment:

The TCEQ has substantially rewritten the DSD so that the main text is focused on the EtO data and the analyses, decisions, and justifications made by the TCEQ in deriving a unit risk factor for EtO. Most of the comparison to the USEPA model has been removed, as well as much of the language criticizing USEPA. In addition, the TCEQ has edited the DSD to include the definition of "supra-linear" as a model with a dose-response curve that is steeper than linear as illustrated in Figure 2 of the final DSD where the low-dose slope is steep beginning at zero dose and then transitions at higher doses to a shallower slope (as is noted by Expert 5). By TCEQ's definition this can include curvilinear models or multi-part linear spline models with this same shape. The TCEQ now largely discusses the USEPA dose-response model as a linear two-piece spline, with a steeper slope in the low-dose region.

Expert 6

Comment 1:

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.

Response to Expert 6 Comment 1:

The TCEQ has edited the DSD to include the definition of "supra-linear" as a model with a dose-response curve that is steeper than linear as illustrated in Figure 2 of the final DSD where the low-dose slope is steep beginning at zero dose and then transitions at higher doses to a shallower slope. By TCEQ's definition this can include curvilinear models or multi-part linear spline models with this same shape. The TCEQ now largely discusses the USEPA dose-response model as a linear two-piece spline, with a steeper slope in the low-dose region.

Comment 2:

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.

Response to Expert 6 Comment 2:

The TCEQ addresses the concerns about the low-dose shape of the dose-response curve in response to Expert 6's comments on charge questions 7 and 8, and the concerns about endogenous EtO levels in response to Expert 6's comments on charge question 10.

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

Comment:

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.

Response to Expert 1 Comment:

The TCEQ appreciates Expert 1's analysis and notes the expert's opinion that the Responses to Public Comments appear to be appropriate and clearly stated.

Expert 2

Comment 1:

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.

Response to Expert 2 Comment 1:

The TCEQ appreciates Expert 2's analysis and notes the expert's opinion that the Responses to Public Comments document is clearly written and addresses the major scientific questions and issues raised by the public commenters.

Comment 2:

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.

Response to Expert 2 Comment 2:

The TCEQ has revised the text in the DSD to more completely describe the use and derivation of the ADAF-adjusted URF. The DSD no longer discusses the *in utero* adjusted factor discussed by the commenters as it refers to a peripherally-related topic not part of TCEQ guidance (TCEQ 2015).

Comment 3:

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.

Response to Expert 2 Comment 3:

Section 4.3.3 of the final DSD provides the scientific and health-protective bases of the TCEQ decision to use the 95% UCL of the slope to derive the URF, and this is consistent with TCEQ guidance (TCEQ 2015).

Comment 4:

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.

Response to Expert 2 Comment 4:

Figure 2 from the RTC has been added to the DSD (Figure 3 in the final DSD) in the context of using MOA data to inform the EtO dose-response assessment.

Comment 5:

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

Response to Expert 2 Comment 5:

The commenter referred to on page 63 of the Response to Public Comments is correct that the Cox proportional hazards model did not significantly fit the data. However, the TCEQ in this response was making the point that none of the models significantly fit the data better than the null model, and choosing to develop a URF despite this lack of significant fit actually represented a health-protective choice because another way to interpret this lack of fit would be to conclude that EtO did not cause cancer in the NIOSH cohort.

Expert 3

Comment:

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

Response to Expert 3 Comment:

The TCEQ appreciates Expert 3's analysis and notes the expert's opinion that the Responses to Public Comments provides a thorough and detailed response to the presented public concerns.

Expert 4

Comment:

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.

Response to Expert 4 Comment:

The TCEQ appreciates Expert 4's analysis and notes the expert's opinion that the Responses to Public Comments appropriately addresses the presented public concerns.

The TCEQ has substantially revised the DSD, and as suggested by this expert, much of the text countering the USEPA assessment is now found in Appendix 6 and in the Responses to Public Comments. The principal focus of the final DSD is on the rationale for the TCEQ EtO hazard and dose-response assessment.

Expert 5

Comment:

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.

Response to Expert 5 Comment:

The TCEQ appreciates Expert 5's comments and addresses specific concerns elsewhere in this document.

Expert 6

Comment:

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 need to be corrected.

Response to Expert 6 Comment:

The TCEQ appreciates Expert 6's comments and addresses specific concerns elsewhere in this document. The TCEQ does not reject the linear two-piece spline model simply because it is referred to as a supra-linear model. For example, the linear two-piece spline model was shown to statistically significantly over-predict the NIOSH lymphoid cancer mortality data, and in the final DSD is now shown to be statistically significantly over-predictive for the UCC cohort as part of a validation study. By contrast, the Cox proportional hazards model based on the NIOSH cohort is statistically shown to be accurate for both the NIOSH cohort and the UCC cohort (validation analysis).

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 3

Comment 1:

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 page 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.

Response to Expert 3 Comment 1:

In Section 4.2.3 of the final DSD the TCEQ focuses on the ability of the Cox proportional hazards and two-piece spline models to predict the NIOSH and UCC cohort data. Most of these analyses are completed with the entire datasets, which should have greater statistical power, and only one analysis considers the predictiveness for specific dose groups. The TCEQ's overall conclusion about model fit is based on all this information together and does not overly rely on any single comparison and whether or not it is statistically significant. The overall data consistently shows that the Cox proportional hazards model produces better predictions of the NIOSH and UCC cohort data, and that the linear two-piece spline model over-predicts the data. The TCEQ has made this conclusion clearer in the final DSD.

Comment 2:

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.

Response to Expert 3 Comment 2:

The TCEQ has substantially revised the DSD, including Chapter 2, to improve organization and logical flow. The information about ecological studies that was included in Chapter 2 has been removed from the final DSD.

Comment 3:

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.

Response to Expert 3 Comment 3:

As suggested by Expert 3, the URF derived by the TCEQ is quite health protective, because it assumes a statistically significant association between EtO dose and lymphoid cancer (while the modeling shows no difference from the null), with use of the 95% upper confidence limit providing further health protection. Furthermore, the URF is based on the more conservative estimate for lymphoid cancer in males (compared to males and females combined) and incorporates USEPA age-dependent adjustment factors, both of which result in a lower, more protective risk-based air concentration for the general public.

These health-protective choices are now more clearly delineated in the DSD.

Expert 4

Comment:

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.

Response to Expert 4:

The TCEQ has revised the DSD to make it clear that the update of the UCC cohort is not used to derive the URF.

Expert 5

Comment:

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.

Response to Expert 5:

The TCEQ appreciates the complexity of these estimates, which are now more thoroughly explained in the DSD and in response to Expert 5's comments on charge question 8.

Expert 6

Comment 1:

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.

Response to Expert 6 Comment 1:

The TCEQ agrees that the best data for predicting low-dose response to low-dose EtO exposure is low-dose data. However, all the occupational and animal study data for EtO are at least thousands of times higher than low-dose (i.e., environmental) exposures. Therefore, TCEQ does not in this case define "lowest available dose" as equivalent to "low-dose" and cannot justify relying exclusively or differentially on the lowest dose occupational data when choosing the shape of the dose-response curve. In the absence of other information, the TCEQ chose to use the putative mutagenic MOA to inform a low-dose linear extrapolation, and we used the standard Cox proportional hazards model for the EtO dose response. One method that TCEQ used to test the models was to assess how accurately the TCEQ and the EPA models predicted the NIOSH cohort data that were used to derive the model. In so doing we found that the TCEQ model can predict the number of lymphoid deaths with 95% confidence overall (Tables 6 and 29) and also in every exposure quintile including the lowest quintile of exposures of the NIOSH study (Tables 7 and 30). In contrast, other proposed models over-predict at the 5% significance level the overall number of lymphoid deaths in the NIOSH study (Tables 6 and 29) and also the number of lymphoid deaths in the lowest quintile of EtO exposures and other quintiles (Tables 7 and 30). This suggests that the TCEQ's model provides the best fit to all the data, including the lowest dose data.

Comment 2:

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.

Response to Expert 6 Comment 2:

Valdez-Flores et al. (2013) conducted an analysis of the EtO NIOSH cohort data using different numbers of categories (4, 20, and 61), and generally found that with an increasing number of categories, the shape of the dose-response changed. Because of this and based on other concerns described in response to charge question 6, the TCEQ does not rely on the categorical modeling results to determine the best-fit model to the NIOSH cohort data. Additionally, the USEPA's SAB indicated that the USEPA should model the individual data, because the individual data provides the best basis for the URF. Categorical data are crude relative to the individual data, which results in the loss of information.

Comment 3:

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.

Response to Expert 6 Comment 3:

The DSD now provides further explanation of the categorical data points in Appendix 6, and this concern is also discussed in response to charge question 6 comments.

Comment 4:

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.

Response to Expert 6 Comment 4:

The TCEQ's hazard assessment of human breast cancer caused by EtO is now explicitly detailed in Section 3.3.1.1 of the DSD. The summary of that section is as follows:

In summary, the epidemiological evidence for EtO causing human breast cancer is very weak, with most of the available studies showing no association when the external reference population is used as a comparison group. This is the same conclusion reached by Marsh et al. (2019) in their recent meta-analysis, which found that there was no evidence from the epidemiology studies of a relationship between EtO exposure and breast cancer. The meta-analysis conducted by Vincent et al. (2019) reached a similar conclusion, stating that "Higher quality epidemiological studies demonstrated no increased risk of breast cancers." When considering the evidence from animal studies, the TCEQ found that while there was an increase in mammary tumors in mice chronically exposed to EtO (NTP 1987), there was no increase in mammary tumors in rats chronically exposed to EtO (Snellings et al. 1984). In addition, IARC in 2019 released an assessment of tumor site concordance, which found that only 20% of the evaluated Group 1 chemicals showed site-concordance of mammary/breast tumors between animals and humans. While the MOA determination that EtO is carcinogenic through mutagenic and genotoxic mechanisms generically supports tumor sites at any location, there is no specific

MOA or metabolic information that identifies breast tissues as a susceptible site. Consistent with the above discussion, the TCEQ determines that there is inadequate evidence for identifying breast cancer as a hazard of EtO exposure in humans.

The TCEQ determined in the hazard assessment that there is inadequate evidence that EtO causes breast cancer, and therefore did not further assess whether the lymphoid cancer assessment was protective of breast cancer.

Reference List

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

Crump, KS, & Allen, BC. 1985. Methods for Quantitative Risk Assessment Using Occupational Studies. The American Statistician 39(4):442-450.

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.

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

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

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

International Agency for Research on Cancer (IARC). 2019. Tumour site concordance and mechanisms in carcinogenesis. IARC Scientific Publication No. 165, Lyon, France: International Agency for Research on Cancer.

Jain, RB. 2020. Associations between observed concentrations of ethylene oxide in whole blood and smoking, exposure to environmental tobacco smoke, and cancers including breast cancer: data for US children, adolescents, and adults. Environmental Science and Pollution Research Available at: https://doi.org/10.1007/s11356-020-08564-z.

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

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.

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.

Marsh, GM, Keeton, KA, Riordan, AS, et al. 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(7):919-939.

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 lymphopoietic 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 B6C3F$_1$ mice exposed to ethylene oxide. Toxicol Sci. 2013 Nov; 136(1):26-38. doi: 10.1093/toxsci/kft190

Recio, L, Donner, M, Abernethy, D, et al. 2004. In vivo mutagenicity and mutation spectrum in the bone marrow and testes of B6C3F1 lacI transgenic mice following inhalation exposure to ethylene oxide. Mutagenesis 19: 215-222.

Rothman, K. J. and J.D. Boice (1979). Epidemiologic Analysis with a Programmable Calculator (NIH Publication 79-1649), Washington, DC, US Government Printing Office

Sielken Jr, R, Valdes-Flores, C. 2009b. Life-table calculations of excess risk for incidence versus mortality. Regulatory Toxicology and Pharmacology. 55:82-89.

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

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 and 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, JA, Fryar-Tita, E, Jeong, YC, Boysen, G, Starr, T, Walker, VE, & Albertini, RJ. 2008. Biomarkers in toxicology and risk assessment: informing critical dose-response relationships. Chem Res Toxicol 21(1):253-265.

TCEQ. 2015. Guidelines to develop toxicity factors. RG-442: Texas Commission on Environmental Quality (TCEQ). https://www.tceq.texas.gov/assets/public/comm_exec/pubs/rg/rg-442.pdf

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

United States Environmental Protection Agency (USEPA). 2005a. Guidelines for Carcinogen Risk Assessment. (EPA/630/P-03/001B). Washington, D.C.

United States Environmental Protection Agency (USEPA). 2012. Benchmark Dose Technical Guidance Document. (EPA/100/R-12/001). Washington, D.C.

United States Environmental Protection Agency (USEPA). 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.

Valdes-Flores, C, Sielken Jr, R, Teta, J. 2010 Quantitative cancer risk assessment based on NIOSH and UCC epidemiological data. Regulatory Toxicology and Pharmacology. 56:312-320.

Valdez-Flores, C, Sielken Jr, R. 2013. Misinterpretation of categorical rate ratios and inappropriate exposure-response model. Regulatory Toxicology and Pharmacology. 67:206-214.

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