June 14, 2023

Submitted via OEHHA Website: https://oehha.ca.gov/comments

Esther Barajas-Ochoa
Office of Environmental Health Hazard Assessment
P. O. Box 4010
Sacramento, California 95812-4010

Re: ACC Comments on Draft No Significant Risk Level (NSRL) for Ethylene Oxide

The Ethylene Oxide Panel of the American Chemistry Council appreciates the opportunity to provide comments on the OEHHA's proposed update to Proposition 65 No Significant Risk Level (NSRL) for ethylene oxide (EtO) and its accompanying Initial Statement of Reasons supporting the update (Proposed NSRL). The Proposed NSRL for EtO is derived by adopting the U.S. Environmental Protection Agency (EPA) Inhalation Cancer Unit Risk (UR) and rejecting the UR derived by the Texas Commission on Environmental Quality (TCEQ) . We strongly urge OEHHA to reconsider this decision and to correct the Proposed NSRL Initial Statement of Reasons of errors impacting exposure-response assessment of the NIOSH lymphoid mortality and breast cancer incidence data.

Revision of the NSRL from 2 µg/day to 0.058 µg/day could have a substantial impact on users of EtO. The IUR from which the NSRL is derived is equivalent to 1 drop of water in a volume of 20 Olympic-sized swimming pools. Potentially impacted uses include ethylene that is metabolized to EtO. Potentially impacted downstream users include: treatment of fruits and vegetables, cosmetics and personal care products, paints and coatings, medical products sterilized with EtO, and numerous other uses. Labeling products that may contain such a minute concentration of EtO will only unnecessarily scare consumers and may result in decreased use of very beneficial products. It is important to note that exogenous and endogenous values of EtO also exceed the NSRL.

EPA's selection of the final Integrated Risk Information System (IRIS) exposure-response model was based on a fundamentally flawed statistical analysis and incorrect assessment of visual fit of categorical (grouped) model estimates. Unfortunately, the OEHHA's uncritical acceptance of

EPA's EtO IRIS 2016 methodology and assumptions results in the same errors and flaws leading to an implausible cancer unit risk (UR) of 6.1 per ppm (3.3×10^{-3} per $\mu g/m^3$) and cancer slope factor (CSF) of 12 per mg/kg-day, from which the Proposed NSRL of 0.058 µg/day was derived. While OEHHA admirably attempts to conduct an "independent" evaluation of bias in EPA's model and the National Institute of Occupational Safety and Health (NIOSH) epidemiological data, their efforts are based on questionable assumptions in the absence of access to the actual data being considered and mimicking EPA's faulty reasoning.

As OEHHA correctly points out, the IRIS UR is based on both lymphoid mortality and breast cancer incidence from the NIOSH study (Steenland et al., 2003, 2004) whereas the TCEQ UR is based on lymphoid mortality alone. As discussed in our detailed comments, the unavailability and incomplete ascertainment of the breast cancer data precludes the use of the NIOSH breast cancer incidence data for quantitative risk assessment purposes. The weaker evidence of causation for breast cancer in the epidemiology studies further supports focus on lymphoid cancer incidence data from the NIOSH study. However, the major reasons for the large difference in magnitude between the two URs are (a) the type of exposure-response model used to fit the epidemiologic data and (b) the interpretation and value of information in related epidemiological and biological evidence to inform selection of the model.

The Description of the NIOSH and Union Carbide Corporation (UCC) cohort studies need to be corrected in the Proposed NSRL Initial Statement of Reasons.

The NIOSH and Union Carbide Corporation (UCC) EtO cohort studies are comparable in terms of the number of lymphoid cancer mortalities (an important factor to consider in the power of the study) and the exposure assessment. While we do not dispute the use of just the NIOSH study to derive the UR/CSF for EtO, the Proposed NSRL description of the strength and weaknesses of these two studies should be corrected so that both the NIOSH and UCC studies can inform selection of the exposure-response model.

Specifically, the Proposed NSRL should limit the description of the NIOSH study exposure assessment as "high quality" and "validated" to apply only to the narrow exposure period after 1978 when data were available to validate the exposure regression model. Prior to 1978 this model is unvalidated because there was very limited or no exposure data available. Furthermore, the model validated after 1978 was altered for years prior to 1978 by holding a key variable "calendar year" fixed at the predicted level in 1978. The calendar year variable is described by the authors of the NIOSH exposure model to be a "surrogate for improvement in work practices." By holding the variable calendar year fixed prior to 1978, the NIOSH exposure model estimates lower exposure estimates in earlier years compared to 1978. No effort was made by the NIOSH authors to independently validate this substantial adjustment to the model prior to 1978.

Bogen et al. (2019) addressed this limitation by finding multiple sources of new information and data indicating changing work practices in earlier years. Ironically, the Proposed NSRL dismisses the Bogen et al. (2019) model showing the opposite but more plausible trend for historical exposures (e.g., higher exposures in earlier years) because *"the authors were unable to validate their pre-1978 predictions since no actual worker measurement were available from that time,"* and thus, *"the accuracy of the Bogen et al. (2019) assessment is unknown to OEHHA."* By the same reasoning, OEHHA must also conclude that the accuracy of the NIOSH model prior to 1978 is unknown, and OEHHA should indicate that the data NIOSH used to develop the model are no longer available (lost) (EPA IRIS, 2016b, Appendix H, p H-28).

Compared to the NIOSH exposure data before 1978, the UCC exposure estimates are superior in quality to the NIOSH exposure estimates because exposure data for the most recent periods, 1957-1973 and 1974-1988, were available for more than 75% of the cohort based on routine monitoring, personal sampling, medical records on severe acute toxicity (e.g. respiratory irritation, nausea and vomiting), and a plant wide survey in another UCC plant using the same process (Greenberg et al., 1990; Teta et al., 1993; Swaen et al., 2009). The UCC cohort experienced more than twice the average estimated cumulative exposure (67 ppm-years) compared with the larger and younger NIOSH cohort (27 ppm-years) (Valdez-Flores et al., 2010). The OEHHA Proposed NSRL Table 1 should be corrected to add the Valdez-Flores et al. (2010) paper and indicate in the third column that 6 different lag years and different exposure scales and models (including log cumulative exposures and cumulative exposures) were analyzed.

These corrections are necessary for a more balanced consideration by OEHHA of both the NIOSH and UCC studies in selecting the most appropriate exposure-response model. The evidence from both of these studies, individually and combined, do not support selection of a 2-slope exposure-response model with a very steep initial slope. The log-linear exposure-response model is far more consistent with the weight of evidence from the NIOSH and UCC studies as described in greater detail in our detailed comments (see Detailed Comments).

Biological and epidemiological evidence should play a primary role in selecting exposure-response model

The exposure-response model used by IRIS (EPA, 2016a) for both breast cancer and lymphoid is a "supralinear"[1] two-slope linear spline model (2-slope model), suggesting that risk increases sharply at low exposures and less steeply at higher cumulative exposures above 1,600 ppm-days for lymphoid cancer and 5,750 ppm-days for breast cancer incidence. In contrast, the

[1] EPA IRIS uses the term "supralinear" to describe the exposure-response relationship. Ironically, EPA (2022) attributes "supralinear" to ACC as if this is not an appropriate description.

dose-response model used by TCEQ (2020a) is the standard[2] log-linear Cox Proportional Hazards (CPH) model, which is virtually linear at relevant exposure concentrations for estimating cancer risk for the general population.[3]

The steep initial slope of the EPA IRIS (2016a) supralinear two-slope spline model, which gives rise to one of the highest inhalation cancer potency estimates derived by IRIS, is not justified based on the relatively weak epidemiological findings reported in the original NIOSH peer-reviewed publications, the weight of evidence in the epidemiological literature including the UCC cohort, and the biological evidence in cancer bioassays and analysis of genotoxicity data.

Our detailed comments explain why the epidemiological evidence and biological evidence is more consistent with the TCEQ model than the IRIS model. While the IRIS assessment includes summaries of the genotoxicity, toxicology, epidemiology and toxicokinetics, there is virtually no integration of these important lines of evidence into the final quantitative risk assessment process. Instead, the IRIS exposure-response assessment is driven by exhaustive statistical modeling analyses divorced from consideration of exposure-response concordance with genotoxicity, toxicology and epidemiological weight-of-evidence. Similarly, OEHHA makes no attempt to integrate the different lines of evidence to inform selection of the exposure-response model. This EPA (2005) carcinogen risk assessment guidelines captures this issue well:

> "Another problem occurs when a multitude of alternatives are presented without sufficient context to make a reasoned judgment about the alternatives. This form of model uncertainty reflects primarily the availability of different computer models and not biological information about the agent being assessed or about carcinogenesis in general. In cases where curve-fitting models are used because the data are not adequate to support a toxicodynamic model, there generally would be no biological basis to choose among alternative curve-fitting models. However, in situations where there are alternative models with significant biological support, the decisionmaker can be informed by the presentation of these alternatives along with their strengths and uncertainties."

Similarly, the EPA SAB (2015) emphasized that "any model that is to be considered reasonable for risk assessment must have a dose-response form that is both biologically plausible and consistent with the observed data." Thus, the epidemiological weight of evidence should play a very important role in the consideration of the model selection. The absence of findings in the UCC study at any exposure, and absence of statistically significant findings at lower exposures in

[2] EPA (2022) attributes the term "standard" to ACC to describe the CPH model TCEQ uses, yet this is the correct term EPA IRIS (2016) uses to describe the log-linear CPH model

[3] EPA (2022) correctly acknowledges on p.57 that "the log linear Cox model is essentially linear in the low dose range".

males in the NIOSH study are more consistent with a standard CPH model than an extremely steep initial exposure-response slope.

In the most recent EPA (2022) response to public comments regarding this lack of consideration of the biological evidence in the dose-response assessment, EPA conducts a highly subjective visual inspection of genotoxicity and cancer bioassay data to support their claim that the biological evidence cannot be used to inform biological plausibility. The EPA (2022) evaluation involved (a) plotting the data as point estimates without error bars, (b) drawing a straight line between the response levels for the lowest and highest dose levels, and (c) declaring the dose-response to be supralinear or sublinear depending on whether the responses for the mid-dose levels visually appeared to be above or below the line. This visual inspection did not involve any consideration of statistical significance or evaluation of which data set and dose regimen is most relevant and useful to inform epidemiology data based on cumulative exposures. Our detailed comments explain why a single slope CPH model linear at lower exposures that gradually increases at higher exposure is much more consistent with the epidemiological, toxicological and genotoxicity evidence compared to a 2-slope model with a very steep initial slope leading to derivation of one of the highest EPA IRIS UR.

The Proposed NSRL model selection criteria is based on flawed statistics and visual fit analysis

The Proposed NSRL reiterates the EPA IRIS (2016a) rationale for selecting the 2-slope model based on statistical and visual fit of different models, without meaningful consideration of important new information that was available in the final TCEQ (2020a) Development Support Document (DSD) and in TCEQ (2020b) response to comments. For example, ***OEHHA reports a p-value of 0.01 for EPA's 2-slope model for breast cancer, suggesting that OEHHA is unaware that independent peer reviewers of the TCEQ DSD who provided in-depth statistics comments agreed with TCEQ that the p-values for the 2-piece spline models were incorrectly calculated.*** The corrected p-values are summarized in Table 1 and indicate that there is no statistical basis to select the 2-slope model over the CPH model. Our detailed comments provide ample evidence for why OEHHA's Proposed NSRL should report the corrected p-values for lymphoid mortality and breast cancer incidence.

Table 1. Corrected p-values for IRIS 2-slope linear spline and IRIS standard CPH model

	EPA IRIS (corrected) 2-slope linear spline	EPA IRIS Standard CPH Model
Lymphoid Mortality	**P=0.14** corrected from 0.07	**P=0.22**
Breast Cancer Incidence	**P=0.04** corrected from 0.01	**P=0.02**

Source: Corrected and IRIS reported p-values are based on IRIS (EPA, 2016a, Tables 4-2, 4-4, 4-6, 4-12, 4-13; EPA, 2016b, Appendix D) and final TCEQ DSD (2020a).

The OEHHA Proposed NSRL perpetuates mistakes with visual fit comparisons made in the IRIS (EPA, 2016a) assessment by stating the following:

> "Other models, including the log-linear models (e.g., Cox regression) and the models using categorical data or exposure transformation, generally resulted in slopes that appear to dramatically over- or under-predict the actual study results, especially in the lower-exposure ranges."

OEHHA makes two errors in this statement:

> **First,** OEHHA incorrectly equates the "actual study results" for lymphoid cancers with the 5 categorical modeled estimates which appear as filled purple dots in IRIS Figure 4-3 with the first "dot" at the origin representing the lagged-out group considered to have zero exposures. Categorical rate ratios (RR) are calculated with respect to a baseline background hazard rate that is also estimated non-parametrically (i.e., not estimated by the CPH procedure). For the continuous models, the actual data modeled are the individual hazard rates not represented graphically in EPA IRIS figures[4]. The true (or implicit) y-intercept (or baseline hazard rate at cumulative exposures) for each continuous model applied to the 53 individual hazard rates will be normalized to 1 at zero lagged exposure.[5] The higher the implicitly modeled y-intercept, the lower the graph will appear on a graph of RR.
>
> **Second,** the OEHHA Proposed NSRL is making the mistake of assessing whether various models underestimate or overestimate the "actual study results" based on the IRIS (EPA, 2016a) subjective visual comparisons along the y-axis. This type of visual comparison is incorrect as clearly indicated in the EPA IRIS (2016a) figure legends: *"Note that, with the exception of the categorical results, the various models have different implicitly estimated baseline risks; thus, they are not strictly comparable to each other in terms of RR values (i.e., along the y-axis)."*[6] In other words, it is impossible to make conclusions about over- or under- predicting the actual study results even if one were to

[4] For lymphoid cancers there are 53 individual hazard rates each reflecting comparisons between one lymphoid mortality case against a risk set of "control" workers defined as those who survived to at least the age of the index case. The risk set of "control" workers includes both exposed and unexposed workers and is distinctly different from the underlying background hazard rate implied by the nonparametric relative rate (or rate ratios; RR) represented by EPA IRIS as the first categorical purple "dot" at the origin.

[5] Allison (2010), a practical guide to survival analysis using SAS, explains that the intercept is part of the arbitrary function of time, which drops out of the estimating equations".

[6] IRIS (EPA, 2016a) Figures 4-2, 4-3, 4-4, 4-5, 4-6, 4-7, 4-9 for lymphoid and breast cancers.

incorrectly define the "actual data" as the 5 categorical RR estimates. As explained above, this is because the baseline background hazard rates implied by the nonparametric (categorical) RRs and the underlying background hazard rates implied by the parametric models are generally different, but when graphed as RR values are all normalized to 1 making it impossible to make any conclusions about under- or over-estimations.

Valdez-Flores and Sielken (2013) and TCEQ (2020a) explain in greater detail why these types of visual comparisons based on IRIS (EPA, 2016a) figures are inappropriate because it cannot be assumed that summary RRs describe the true underlying exposure–response relationship for the continuous models. The SAB (2007) implied the same recommendation against visual fit when they instructed EPA to use the individual data to fit the dose-response models. The SAB (2007) concluded "The Panel was unanimous in its recommendation that the EPA develop its risk models based on direct analysis of the individual exposure and cancer outcome data for the NIOSH cohort rather than the approach based on grouped data that is presently used."

OEHHA Proposed NSRL should reconsider TCEQ's objective method for assessing model fit based on theoretical concerns that were addressed, instead of relying on incorrect statistical and visual fit methods

Compared to the IRIS and OEHHA visual "eyeballing" comparisons using graphs that are not fit for this purpose, TCEQ (2020a) provides a far more objective method to check how well each of the two models (i.e., the standard CPH model[7] and the 2-piece spline model[8]) applied to general population background cancer rates can predict the number of lymphoid cancer deaths (the key cancer endpoint) that were actually observed in the NIOSH cohort. This approach used to predict cancer deaths from the model is essentially the same well-accepted approach IRIS (2016a, Section 4.7) used to estimate extra risk for various occupational exposure levels by applying the IRIS model to the general population background cancer rates in the life-table program.

Table 2 compares the number of lymphoid cancer deaths that were observed in the NIOSH cohort versus the predictions by the IRIS and TCEQ models. In this model ground-truthing exercise, the TCEQ model was not only able to better predict the actual total number of lymphoid cancers in the NIOSH cohort but also the number of cancers in Quintile 2[9] demonstrating superior global and local fit below the knot.

[7] TCEQ's model using 15-yr lag and the full risk set.

[8] EPA IRIS model using 15-yr lag and 100 workers randomly selected individuals from each case's risk set.

[9] TCEQ (2020a) defined Quintile 1 as the 9 lagged-out cases (no exposures). The remaining 44 cases were equally divided into 4 groups designated by TCEQ as Quintiles 2-5. EPA IRIS (2016a, p. 4-15) reported 13 exposed cases below the knot of 1,600 ppm-days. Thus, prediction of Quintile 2 comprised of 11 cases with the lowest exposures best reflects "local" fit below the knot.

Table 2. Comparison of observed versus predicted number of lymphoid cancer deaths in NIOSH study using IRIS and TCEQ models

	Observed	Predicted (95% CI)	
		2-piece linear spline model with knot	Standard log-linear CPH model (linear at POD of 1/100 and below)
Total number of cancer deaths	53	91.69 (70.1, 122.4)	52.42 (40.1, 70.0)
Quintile 2[8] cancer deaths	11	20.9 (11.7, 42.0)	14.4 (8.1, 28.9)

Source: TCEQ (2020a) Appendix 3 Table 29 and 30.

OEHHA's dismissed TCEQ's "Reality Check" using a prediction method based on an unsupported vague claim that the analysis is flawed because "TCEQ's calculations did not accurately account for any differences that might exist between the general US population and the NIOSH worker cohort." The epidemiologic literature has shown that a healthy worker effect (HWE) is predominately related to workers with shorter follow-up and non-cancer causes (Monson, 1986; Fox and Collier, 1976). Most importantly, the NIOSH study authors themselves concluded that there was unlikely to be a cancer HWE in this longer follow-up study (Steenland et al., 2004[10]). Their conclusions are consistent with that of the International Agency for Research on Cancer (IARC) textbook, *Cancer Epidemiology: Principles and Methods* (IARC, 1999), which specifically notes that HWE "is known to vary with type of disease, being smaller for cancer than for other major diseases, and it tends to disappear with time since recruitment into the workforce." Thus, it is unlikely that there is a HWE for the cancers of interest in the NIOSH cohort.

OEHHA did not appear to be aware of the TCEQ DSD[11] sensitivity analysis to demonstrate that the TCEQ model still better predicts the overall actual cancers after applying a high HWE of 15-16% for lymphoid cancers. Although one can quibble with the TCEQ's selection of 15-16% based on ***overall***[12] cancer SMRs from a Norwegian worker study with relatively short average follow-up of 11.5 yrs (Kirkeleit et al., 2013), the larger point is that 15% HWE is a reasonable estimate for differences between the general population and the NIOSH worker cohort given the

[10] "The healthy worker effect would seem an unlikely explanation for the lack of cancer excesses in the exposed versus non-exposed comparisons." (Steenland et al., 2004)

[11] TCEQ (2020a) Section A3.3.2.

[12] Kirkeleit et al. (2013) did not find a HWE for lymphoid (SMR of 0.97 for males, 1.07 for females) or breast cancer (SMR of 1.02), but TCEQ conservatively used the overall cancer SMR of 0.85 and 0.84 for male and female workers. It is also unknown whether the Norwegian workers are representative of the NIOSH sterilizer workers.

unlikelihood of a cancer HWE in the NIOSH study (Steenland et al., 2004). Figures 1 and 2 show the TCEQ model better predicts the observed lymphoid deaths than the IRIS model overall, and locally at Quintile 2. **Taken together, OEHHA's Proposed NSRL should be corrected to indicate that the TCEQ model has excellent overall and local fit based on the TCEQ's prediction analysis, which considers a possible HWE effect as a reasonable surrogate for differences that might exist between the general US population and the NIOSH worker cohort.**

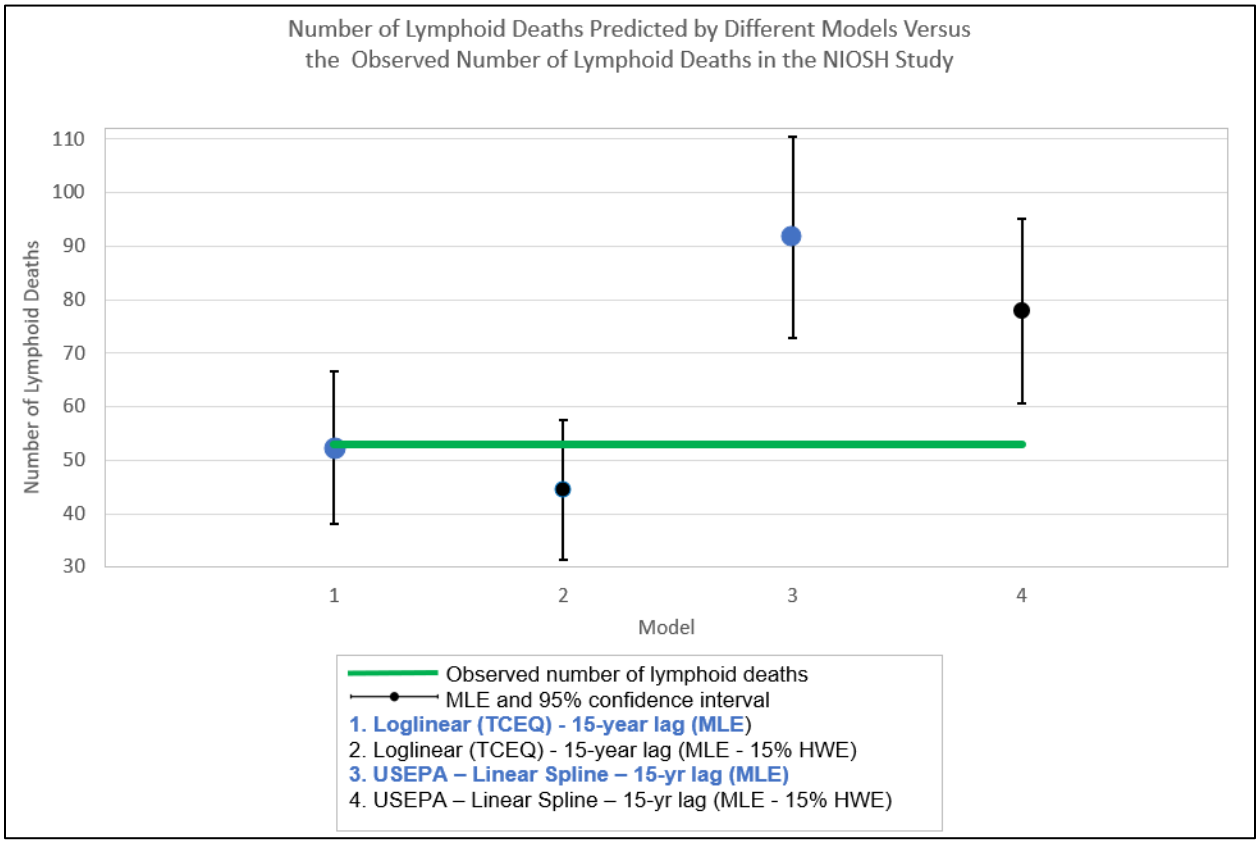

Figure 1. Comparison of <u>overall</u> tit of TCEQ and IRIS models to the NIOSH study

Note: Confidence intervals (CI's) shown are based on Poisson distribution and are very similar to those calculated by TCEQ (2020a) as described in our detailed comments. These similar results provide additional support for TCEQ's conclusions that the TCEQ model has greater overall fit to the data. (See TCEQ, 2020a)

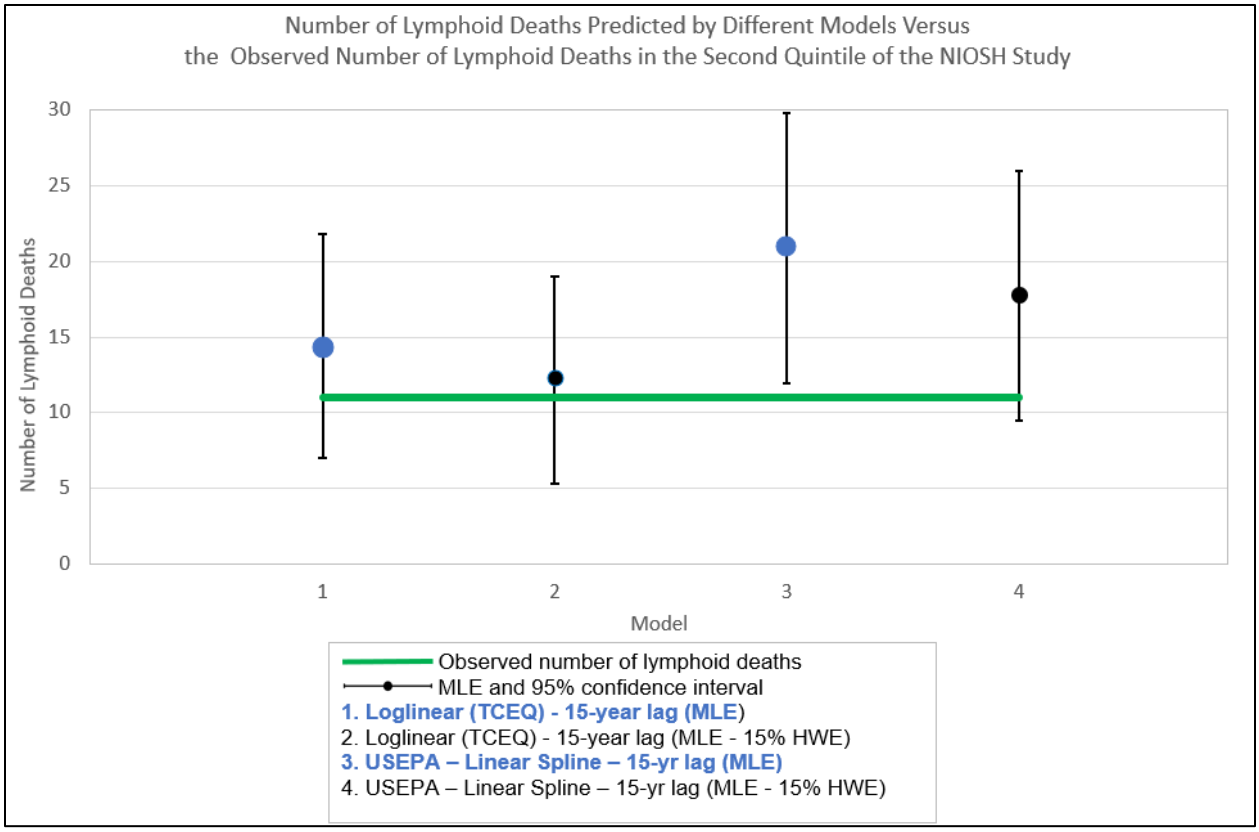

Figure 2. Comparison of <u>local</u> fit of TCEQ and IRIS models to Quintile 2 of the NIOSH study
Note: CI's shown are based on Poisson distribution and are very similar to those calculated by TCEQ (2020a) as described in our detailed comments. These similar results provide additional support for TCEQ's conclusions that the TCEQ model has superior *local* fit below the knot.

Background endogenous and ambient levels of EtO provide important reality checks for model selection. The key assumptions in extrapolating the exposure-response relationships at lower exposures are scientifically valid and are now corroborated by independent data sets.

Background endogenous and ambient levels of EtO are an important reality check for the TCEQ and IRIS model. While OEHHA is correct that the potency estimates technically only apply to exposures above endogenous levels, it is implausible that a chemical would be a potent carcinogen at levels that the body produces through natural processes and also well within the population variability.

We applaud OEHHA for including a section on endogenous levels. However, we disagree with OEHHA's conclusion that the exposure-response relationship for endogenous ethylene oxide exposures is unknown, and therefore cannot be estimated using the Kirman et al. (2021) method. The linear relationship between hemoglobin N-2-hydroxyethylvaline (HEV) adducts

and EtO exposures is well-supported by data across a broad range of exposure levels (i.e., ranging from background levels to ~4 ppm in workers). As explained in detailed comments below, this well-supported linear relationship is also supported by a "forward" analysis suggested by EPA (2022) based on measured EtO concentrations in mainstream cigarette smoke that corroborates the Kirman et al. (2021) linear relationship.

In summary, we urge OEHHA to adopt the TCEQ model for lymphoid cancers or use CPH model estimates from the IRIS assessment and revise the Proposed NSRL to better reflect the following:

1. **The original NIOSH study upon which OEHHA's Proposed NSRL is based found no indication of increased risk of lymphoid cancers in males at lower categories of exposures and concluded there were no exposure-related effects in females. The TCEQ dose response model is more consistent with the original findings of the NIOSH mortality study.**

2. **Breast cancer, like other types of cancers OEHHA considered from both animal and human studies, is a cancer endpoint that deserves consideration in the weight of evidence for cancer classification. However, the NIOSH breast cancer incidence data should not be used for <u>quantitative</u> risk assessment based on substantial under-ascertainment of incident cases reported by Steenland et al. (2003) and subsequent risk deficits in the lower exposures.**

3. **OEHHA's rationale for supporting EPA IRIS model selection is based on uncritical acceptance of EPA's incorrect statistical analysis that did not account for EPA's systematic statistical search for the knot as an estimated statistical parameter. Independent peer reviewers for TCEQ agreed with TCEQ's corrections of the statistics, which puts the TCEQ model on par with the IRIS model based on statistical significance alone.**

4. **The TCEQ exposure-response model is much more plausible based on the biological and toxicological evidence, and the mode of action.**

5. **OEHHA appears to equate the 5 categorical rate ratios with the 53 rate ratios, and/or assumes that comparisons of the exposure-response curve can be compared visually with the EPA IRIS (2016a) graphical representation 5 categorical rate ratios.**

6. **OEHHA dismisses TCEQ's prediction analysis which is a much more objective method than visual fit to check how well each of the two models (i.e., TCEQ's vs EPA's) can predict the observed number of lymphoid cancer deaths (the key cancer endpoint). In this model ground-truthing exercise, the TCEQ model better predicted not only the overall number**

of lymphoid cancers in the NIOSH cohort but also the observed cancers below the knot in the lowest exposure quintile.[13]

7. **OEHHA's sole reliance on internal analyses and OEHHA's extreme and complete exclusion of external analysis is based on a main conclusion that all external analysis should be ignored because it is confounded by HWE. This is contradicted by NIOSH study authors' own published conclusions that "the healthy worker effect would seem an unlikely explanation for the lack of cancer excesses in the exposed versus non-exposed comparisons." Furthermore, OEHHA's uncritical acceptance of conclusions of a more recent paper by Park (2020) that there is a Healthy Worker Survival Effect (HWSE) led OEHHA to support EPA IRIS model. However, the conclusions are not supported by the actual results in the paper. Improvements in these and other descriptions of the epidemiological evidence are needed to accurately assess the epidemiological weight-of-evidence.**

8. **Based on uncritical acceptance of the IRIS evaluation, OEHHA inaccurately exaggerates the reliability of the NIOSH worker exposure estimates prior to 1978.**

9. **The Union Carbide Corporation (UCC) cohort should play a prominent role in considering the strength and consistency of the epidemiology data in supporting the IRIS vs. TCEQ UR. OEHHA's description of this cohort incorrectly omits the internal analysis by Valdez-Flores et al. (2010) which included exploration of different exposure metrics and lag times. The absence of findings in the UCC cohort for male lymphoid cancer mortality is not consistent with a steep slope at low concentrations.**

10. **While EPA's potency estimate technically only applies to exposures above endogenous levels, it is implausible that a chemical would be a potent carcinogen at levels that the body produces through natural processes. The key assumptions in extrapolating the dose-response relationships at lower exposures are scientifically valid and are now corroborated by independent data sets on smoking.**

11. **OEHHA cites two studies reporting an association between smoking and lymphohematopoietic (LH) cancers published in 2012 (Diver et al. 2012 and Kroll et al. 2012) to discount Kirman et al. (2021) reality checks. These two studies are inconsistent with the lymphoid cancer (NHL, lymphocytic leukemia, multiple myeloma) findings from the NIOSH mortality study (Steenland et al., 2004), upon which IRIS 2016 developed their**

[13] TCEQ (2020a) defined Quintile 1 as the 9 lagged-out cases (no exposures). The remaining 44 cases were equally divided into 4 groups designated by TCEQ as Quintiles 2-5. EPA IRIS (2016, p. 4-15) reported 13 exposed cases below the knot of 1600 ppm-days. Thus, prediction of Quintile 2 comprised of 11 cases with the lowest exposures best reflects "local" fit below the knot.

low exposure high risk model and do not constitute a weight-of-evidence evaluation akin to the Surgeon General report, *The Health Consequences of Smoking —50 Years of Progress* (US DHHS 2014). In addition, OEHHA cites IARC review indicating a positive association between tobacco smoking and breast cancer, though not for lymphoid cancer.

12. **OEHHA's Proposed NSRL is an estimated EtO 10^{-5} risk specific intake level. This risk-specific intake level provides little utility in managing general population risk if background exogenous exposure isn't considered as an initial reality check. There has been an extensive ambient air measurement campaign over the last several years, including measurements near many sterilizer facilities and at background locations. In many cases, the levels of ethylene oxide far away from sterilizer facilities are similar to the levels near sterilizer facilities. Although the sources makeup of this exogenous background ethylene oxide is currently not fully characterized, what is clear is that, in many cases, residents living near sterilizer facilities are not exposed to higher ethylene oxide than people living far away.**

DETAILED COMMENTS

1. The original NIOSH study upon which OEHHA's Proposed NSRL is based found no indication of increased risk of lymphoid cancers in males at lower categories of exposures and concluded there were no exposure-related effects in females. The TCEQ dose response model is more consistent with the original findings of the NIOSH mortality study.

** Pertaining to NSRL p. 23-38 **

The extremely steep dose response model selected by EPA IRIS (2016a) and adopted by OEHHA for lymphoid cancer mortality is based on an UR that is among the highest EPA IRIS inhalation URs for known or likely carcinogens. This is inconsistent with the following original conclusions by the NIOSH study authors regarding both internal and external comparisons (Steenland et al., 2004):

- "There was little evidence of any excess cancer mortality for the cohort as a whole"
- "The healthy worker effect would seem an unlikely explanation for the lack of cancer excesses in the exposed versus non-exposed comparisons"
- "Positive exposure-response trends for lymphoid tumors were found for males only (15-year lag)"
- "It is not known why we find an association for males and not females for haematopoietic cancer. . .there was sufficient variation in the exposure of women to have observed an exposure-response if one existed.

A large number of models were considered by IRIS, including those using log transformation of cumulative exposure, which IRIS (EPA, 2016a) correctly excluded as "biologically implausible." Of the models using cumulative exposures, the strongest trend was seen in male lymphoid mortality. As described in detail in the next section, breast cancer incidence is not an appropriate endpoint based on the weight-of-evidence and quality issues. Therefore, of the critical endpoints selected by IRIS, male lymphoid mortality is the most appropriate endpoint for risk assessment, protective of effects in females who showed no sensitivity.

In the internal categorical analysis for male lymphoid tumors, the only statistically significant increase in male lymphoid odds ratio (OR) is only at the highest cumulative exposure level. Steenland et al. (2004) concluded that there was no association of EtO with lymphoid cancers in females. These peer-reviewed conclusions of the original authors are far more consistent with the shallower linear slope of the TCEQ's log-linear exposure response model as compared to a steep initial exposure response of the EPA's 2-slope model.

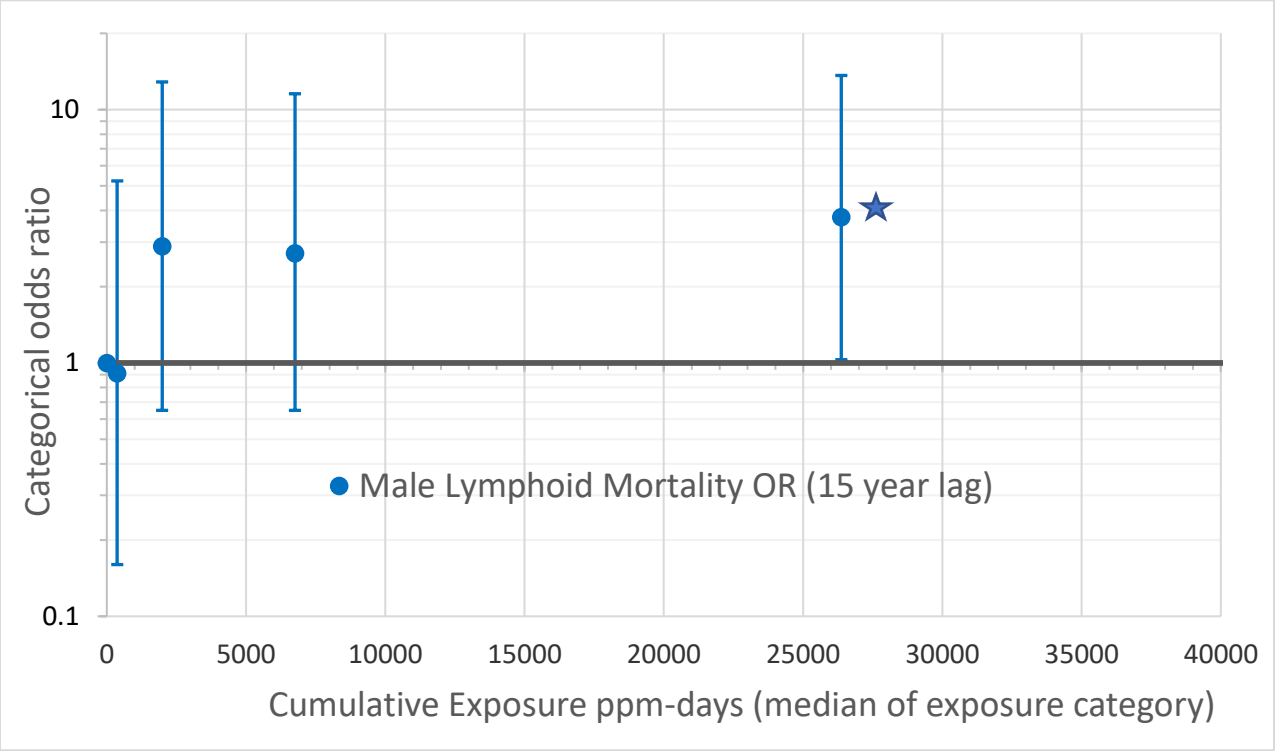

Figure 3. Categorical Odds ratios (OR) for males (95% Confidence Intervals):
Note: Male ORs for exposure categories 0 (lagged out), >0-1,200 ppm-days, 1,201-3,680 ppm-days, 3,681-13,500 ppm-days, >13,400 ppm-days are, respectively, 1, 0.91 (0.16-5.23), 2.89 (0.65-12.86), 2.71 (0.65-11.55), **3.76 (1.03-13.64**) (EPA IRIS, 2016b, Table D-28)

Statistical modeling can take on a life of its own if not checked against the epidemiological data which do not indicate an extremely steep low-dose dose response and/or a major discontinuity in the dose response. The UCC study of EtO chemical workers with comparable numbers of lymphoid cancers and substantial exposures to EtO show no increased risk of male lymphoid cancers. It is important to keep into perspective that the relevant epidemiology, including a large number of human studies published over a forty-year period, indicates that there is inconclusive evidence of carcinogenicity (IARC 2012a). Taken together, the findings from EO epidemiology conflict with the IRIS risk values which imply EtO is a highly potent carcinogen at lower cumulative exposures.

2. Breast cancer, like other types of cancers OEHHA considered from both animal and human studies, is a cancer endpoint that deserves consideration in the weight of evidence for cancer classification. However, the NIOSH breast cancer incidence data should not be used for <u>quantitative</u> risk assessment based on substantial under-ascertainment of incident cases reported by Steenland et al. (2003) and subsequent risk deficits in the lower exposures.

** Pertaining to NSRL p. 29, 35 **

The primary reason breast cancer should not be included in EtO quantitative exposure response analysis is that there is a substantial under-ascertainment of incident cases due to non-participation in the interview study that raises a serious potential selection bias. In addition, the evidence related to causation is weak for breast cancer, supporting the focus on lymphoid cancers for quantitative risk assessment based on the epidemiological data.

Neither the NIOSH breast cancer incidence study (Steenland et al., 2003) nor the NIOSH mortality study (Steenland et al., 2004) report an overall excess of breast cancer. The positive NIOSH findings based on internal analyses are not robust in that they are seen with a certain lag and exposure metric that are not evident with numerous other exposure metrics, models, or lags. The breast cancer incidence findings are at most suggestive, not only due to inconsistencies in the exposure-response, but also due to incomplete cancer ascertainment and the subsequent potential for bias. This disease endpoint is only weakly supported by other epidemiology studies and is inconsistent with others. Thus, the NIOSH study breast cancer incidence findings are not consistent with the selection of a 2-slope model with a steep initial slope.

The IRIS breast cancer incidence analysis relied on data from the subpopulation of the NIOSH cohort that was interviewed, which required both locating subjects and identifying those diagnosed with breast cancer. Of the 7,576 women in the NIOSH cohort, only 5,139 (68%) were included in the interview portion of the study. The percent non-response was of concern, according to the authors. The majority of these, 22%, could not be located and therefore any breast cancer diagnosis would have been missed. Steenland et al. (2003) indicated that cases lost are more likely to be shorter term (i.e., lower cumulative exposure) employees. Those who work longer (i.e., higher cumulative exposures) stay in the area longer and are more likely to get picked up in the state tumor registries and be found for interview. Shorter duration workers with lower cumulative exposures are more likely to leave the area and not be captured in the overall analyses and less likely to be interviewed. If more cases were missed among those with lower cumulative exposures (shorter term employees), then the data would be biased toward seeing a positive slope and/or elevated risk in the higher exposure groups, as reported by Steenland et al (2003).

Steenland et al. (2003) stated that "breast cancer ascertainment in the sub-cohort with interviews was considered complete." In other words, all the women who were interviewed were identified as having breast cancer or not. This, however, does not account for the missing cases among non-participants in the interview study or for cases never identified in the overall target population. Importantly, there is no way of knowing that the distribution of cases by level of exposure in the subcohort of interviewed breast cancer cases is comparable to the distribution in the fully ascertained total cohort. Due to the greater difficulty of locating women with short term employment, there is a high potential for bias in missing cases at lower cumulative exposure. The rate ratios for breast cancer incidence in the lowest exposed groups in the entire cohort were 0.88 (15 yr. lagged out group) and 0.74 (<647 ppm-days, no lag), the latter of which was a statistically significant deficit (Steenland et al., 2003, Table 3). These deficits contributed to suggested positive trends.

Steenland et al. (2003) made an attempt to investigate possible selection bias but noted he didn't have adequate data to address this concern:

> "A second possible bias was the preferential ascertainment of breast cancer among women with stable residence in states with cancer registries; women with stable residency might be expected to have longer duration of employment in companies under study, and hence greater cumulative exposure. Unfortunately, we didn't have residential history, limiting our ability to explore this possibility." (Steenland et al., 2003)

Without the incidence data, selection bias cannot be properly tested. However, several issues support this explanation: 1) the overall population long-term (higher exposed) women would be easier to identify as having breast cancer as they remain for longer periods in states of employment that have tumor registries; 2) similarly, longer employed and higher cumulative exposure women would be easier to locate and thus interviewed; and 3) duration of employment in Steenland et al. 2003 showed a stronger relationship with breast cancer than did actual cumulative exposure.

Steenland et al. (2003) recognized this limitation as one reason the authors were tempered in their causation conclusions:

> "Our data suggest that ETO is associated with breast cancer, but a causal interpretation is weakened due to some inconsistencies in exposure-response trends and possible biases due to non-response and incomplete cancer ascertainment."

In summary, the weak association of EtO with breast cancer and the exposure-response uncertainty due to the sizeable number of missing breast cancer cases precludes use of the NIOSH interview data in derivation of UR or CSF, and the subsequent NSRL. This, together with

the unavailability of the breast cancer incidence data to other researchers to independently examine these issues raises quality issues that indicate the data are inappropriate for exposure-response modeling for regulatory cancer risk assessment purposes.

3. OEHHA's rationale for supporting EPA IRIS model selection is based on uncritical acceptance of EPA's incorrect statistical analysis that did not account for EPA's systematic statistical search for the knot as an estimated parameter. Independent peer reviewers for TCEQ agreed with TCEQ's corrections of the statistics, which puts the TCEQ model on par with the IRIS model based on statistical significance alone.

** Pertaining to NSRL p. 26, 37 **

OEHHA uncritically accepts EPA IRIS (2016a) conclusions on model fit that are based on incorrect statistical analysis and inappropriate visual fit comparisons (described below). TCEQ (2020a) provided corrected Akaike information criterion (AIC) and p-values for the spline models that OEHHA could have easily verified were correct. The peer review of TCEQ dose-response assessment included two independent reviewers who provided in-depth statistical review. Both peer reviewers agreed that EPA incorrectly calculated p-values because they did NOT correct for including the knot as an estimated parameter in the model, a basic violation of statistical principles:

> **Expert 5:** "I do believe that TCEQ has identified a real problem with the USEPA AIC and p-value calculations. The explanation of the issue and the resolution supplied in DSD seems appropriate. That is, I agree with TCEQ that the knot parameter in the spline models should be considered in the count of the parameters, that the AICs reported by USEPA for those models are too low by a value of 2, and that the p-values should be computed using an approximation to a chi-square with 3 degrees of freedom." (TCEQ 2020b, p. 45)
>
> **Expert 6:** "I consider that the location of the spline should be considered a parameter when evaluating fits of spline models, as long as the data were used in determining the knot, as it apparently was in EPA's model. I believe also that the lag should also be considered a parameter when the data are used to determine its value. But, in general, I consider the AIC in such complex models to be essentially only a rough guide to evaluating fit. Therefore, I think TCEQ's conclusion that the 'lower AIC means that TCEQ's selected model is a statistically superior model fit than USEPA's selected model when taking into account model complexity' is an overstatement. Comparing a model with an AIC = 464.5 to one with an AIC = 264.4[sic[14]], 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." (TCEQ 2020b, p. 50)

[14] Expert really meant 464.4

The basic principle of accounting for all modeled parameter is clearly articulated in the National Research Council report entitled "Models in Environmental Regulatory Decision Making", which states that the strategy to pick the "best model" for regulatory decision making should be "subject to a penalty function reflecting the number of model parameters, thus effectively forcing a trade-off between improving model fit by adding addition[al estimated] model parameters versus having a parsimonious description" (NRC, 2007, pp. 174). Importantly, there are no recognized exceptions to the penalty component of the balance incorporated into the AIC metric when applied in a valid procedure for model-selection (Burnham et al., 2002). This general principle is well recognized in the peer reviewed literature to apply specifically to including the estimated "knot" or inflection point from reporting the use of 2-piece linear spline models (Berman et al., 1996; Li et al., 2011; Gkioulekas et al., 2018; Molinari et al., 2001).

EPA Office of Research and Development (ORD) incorrectly claims that they fixed the knot and then conducted a sensitivity analysis. However, EPA ORD did not simply "fix" or "select" the position of the knot in that model. Instead, IRIS systematically tested 20 alternative knots for breast cancer and 70 knots for lymphoid mortality, and then among these, selected knot values that maximized the likelihood of data fit to a corresponding 2-piece spline model.

In the EPA IRIS (2016b, Appendix D at p. D-13), Dr. Steenland provided statistical analysis considering the knot as a parameter for breast cancer to show this had no substantial effect in that analysis, but a similar examination was not presented in the case of lymphoid cancer. In other words, there was clear acknowledgement and recognition expressed in the IRIS (EPA, 2016b) assessment that each knot value that was used to obtain a final spline-model fit is appropriately interpreted as an estimated parameter. Thus, IRIS should have reported the p-values considering the knot as an estimated parameter for breast and lymphoid cancers in the summary tables of the main report for greater transparency.

OEHHA should revise their discussion of fit of the data to reflect the corrected p-values reported by TCEQ (2020a) for the IRIS selected 2-piece spline provided in Table 1 of our comments above. Based on corrected statistical analysis alone, neither EPA IRIS (p=0.14) nor the TCEQ (p=0.22) exposure-response models for lymphoid cancers provide strong evidence that the exposure-response slope differs from zero. Based on statistics alone, the CPH model fits the data similarly to the supralinear 2-piece spline slope but has the advantage of parsimony (simpler model) and biological plausibility (described below). Also described in greater detail below, the CPH model more accurately predicts the observed lymphoid mortalities overall and at lower exposures in the NIOSH study compared to the IRIS (EPA, 2016a) selected 2-piece spline model.

In addition to Table 1, we suggest OEHHA include Table 3, below, in its detailed analyses, which provides more complete and direct comparison between the statistics and UR derivation for the 2-piece spline model and the CPH model. Table 3 will provide users of OEHHA's assessment an

understanding of the range of values that can be estimated for risk assessment based on the IRIS (EPA, 2016a) preferred methods and assumptions.

Table 3. Comparison of IRIS (EPA, 2016a) derived models (URs not including the ADAF)

	2-piece linear spline Linear at EPA POD 1/100	**Standard CPH Linear at EPA POD 1/100**
Model of individual data?	Yes	Yes
IRIS full model name	Linear spline model with knot at 1,600 ppm x days	Log-linear model (standard Cox regression model)
LYMPHOID INCIDENCE (Males and Females)		
IRIS p-value	**0.14 corrected from 0.07**	**0.22**
Central estimate UR (per ppm)	1.34	0.0095
Upper bound UR (per ppm)	5.26	0.020
Note: IRIS (EPA, 2016a) derived UR for lymphoid incidence data based on slope in the fitted mortality model together with background incidence rates in a life-table calculation. It is incorrect to assume that a cancer slope based on mortality can be applied to incidence data (see Sielken and Valdez-Flores, 2009 for detailed explanation)		
BREAST CANCER INCIDENCE (Females)		
IRIS p-value	**0.04 corrected from 0.01**	**0.02**
Central estimate UR (per ppm)	0.71	0.08
Upper bound UR (per ppm)	1.48	0.14
These data are not appropriate for quantitative risk assessment purposes because authors report substantial number of missing cases with higher potential for those with shorter employment missing (Steenland et al., 2003). These data have not been available for independent evaluation by EPA or the public, and, thus, lack transparency, verification, and independent analysis.		
LYMPHOID & BREAST CANCER INCIDENCE (Males and Females)		
Central estimate UR (per ppm)	2.1	0.1
Upper bound UR (per ppm)	6.1	0.15
These data are not appropriate for risk assessment because the breast cancer incidence data are included. EPA provided no justification for the POD of 1/100. TCEQ (2020a) analysis shows that the POD 1/100 for the standard CPH model extrapolates above or in the high range of the experimental data!		

4. The TCEQ exposure-response model is much more plausible based on the biological and toxicological evidence, and the mode of action.

** Pertaining to NSRL p. 15-18, 20-38**

A biological mode of action should be a major consideration when selecting a model for risk assessment.

While there is clear evidence that EtO is genotoxic and carcinogenic, this does not necessarily mean that EtO is acting through a genotoxic MoA for its carcinogenicity. Currently, this MoA should be considered as a default assumption in the absence of a convincing alternate MoA that does not involve genotoxicity as the initial key event. The dose-response and temporality of EtO induced genotoxicity in the etiology of either animal or human tumors has not been fully vetted through a formal process such as the one recommended by International Program on Chemical Safety (Boobis et al., 2006).

For a direct acting alkylating agent such as EtO, the default dose-response for the induction of mutations is linear. This is the worst-case scenario since at low doses closer to the origin, one should expect cellular protective mechanisms (e.g., detoxification and DNA repair) to offer protection, resulting in a shallower slope in this region when compared to higher doses. Based on a presumed genotoxic MoA, both TCEQ and EPA/OEEHA estimate cancer risk based on a linear extrapolation from the POD to the origin but apply very different statistical models to the same epidemiological study to derive the POD, i.e., Cox proportional hazards (CPH) model by TCEQ vs. the two-piece spline model by the EPA/OEEHA. In the 2-two-piece spline model, the initial slope rises rapidly at lower exposure levels and then rises more gradually for higher exposures. This type of dose-response is not consistent with the biology of how EtO works as a direct acting genotoxicant. The dose-response curve for EtO-induced gene mutations in the bone marrow (Recio et al., 2004, Figure 4) and lung (Manjanatha et al., 2017, Figure 5) tissues of transgenic Big Blue mice is especially informative since these tissues represent targets for EtO-induced tumors. In both cases, there is no evidence for a steeper initial slope and the dose-response pattern is more consistent with the CPH model than the 2-piece spline.

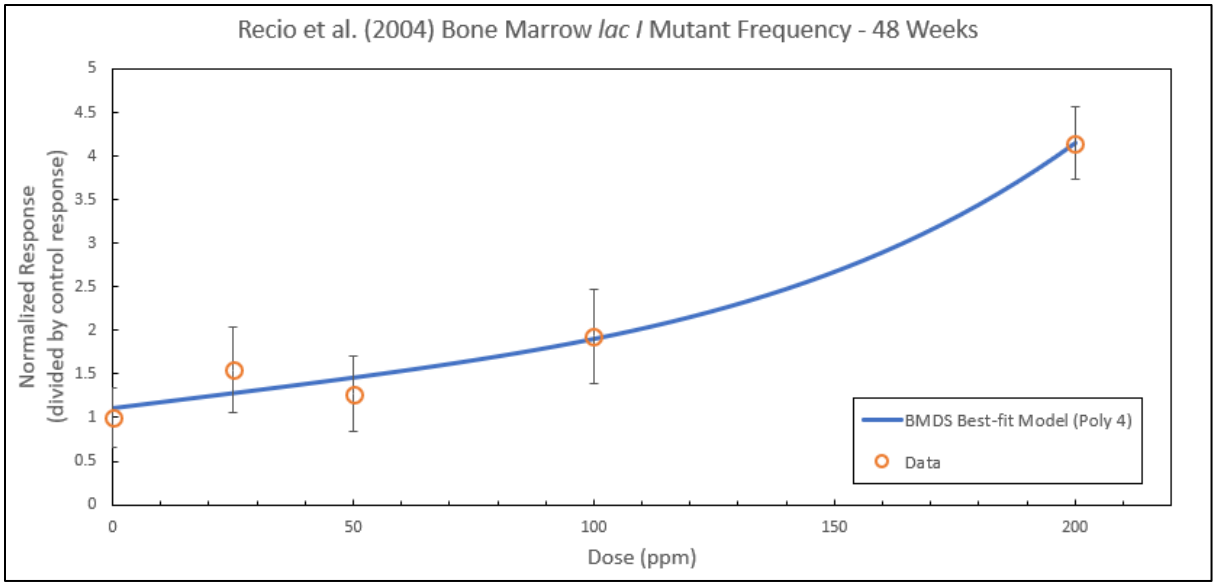

Figure 4. Dose-Response for EtO-Induced *lac*I mutations in Mouse Bone Marrow (6 h/day; 5days/week; 48 weeks from Recio et al., (2004)

Recio et al. (2004) observed increases in *lac*I mutant frequency in the bone marrow of transgenic Big Blue B6C3F1 mice at EtO exposure concentrations of 100 and 200 ppm (but not at 25 or 50 ppm) after 48 weeks of exposure. No increases were observed following 12 and 24 weeks of exposure.

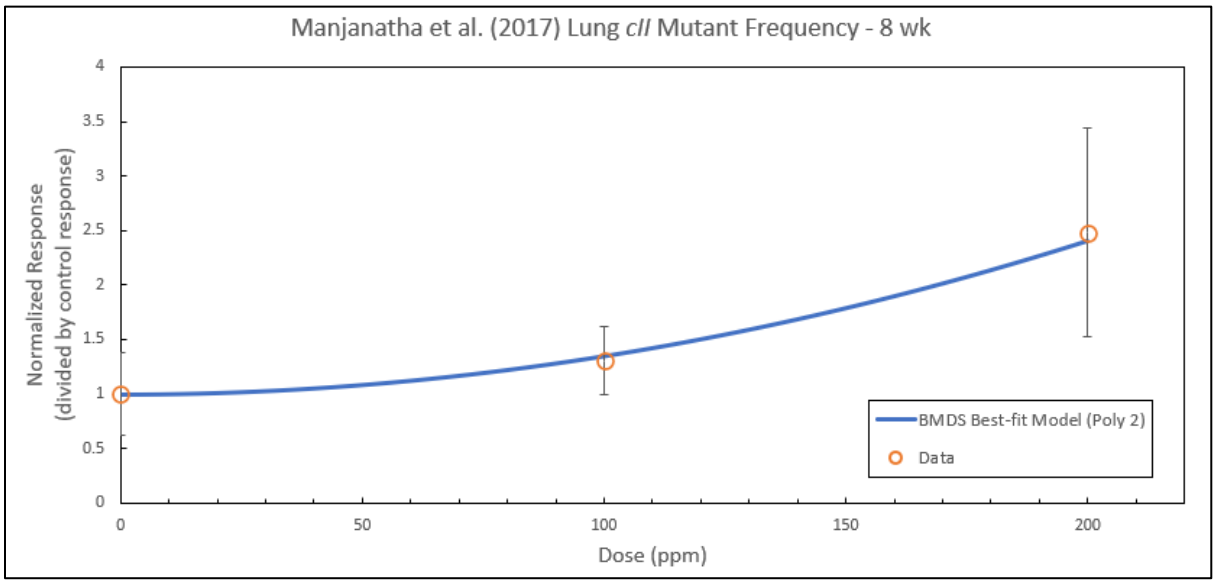

Figure 5. Dose-response for the Induction of *cII* mutations in the lung tissue of Big Blue B6C3F1 transgenic mice at 8 weeks of inhalation exposure to EtO (Manjanatha et al., 2017).

Similarly, Manjanatha et al. 22 (2017) investigated exposure-response and temporality for EtO-induced mutations at the *cII* locus in the lung tissue of transgenic Big Blue male B6C3F1 mice exposed to 0, 10, 50, 100, or 200 ppm EtO for 6 hr/day, 5 days/week over 4 weeks (in all

exposure groups) or for 8 or 12 weeks (in only the two highest exposure groups). A significant increase was observed only following 8 or 12 weeks of exposure and only at the highest concentration studied (200 ppm), which was twice the tumorigenic concentration used in the NTP (1987) bioassay in the same strain of animal. Contrary to expectations consistent with a mutagenic mode-of-action (MOA), no statistically significant increase in mutant frequency or mutational spectrum were observed following 4 weeks of EtO exposure (which is considered to be adequate exposure duration for detecting chemically-induced mutations as per OECD test guideline 488). These results are inconsistent with modified Hill criteria for exposure-response and temporality assuming a mutagenic MOA when considering the NTP (1987) studies in male and female B6C3F1 mice exposed to 0, 50, 100 ppm, 6 hrs/day, 5 days/wk for 102 weeks.

The above dose-response patterns are fully consistent with the molecular initiating event(s) leading to EtO-induced mutagenicity, i.e., the formation for DNA adducts. Marsden et al. (2009) using a highly sensitive liquid chromatography-tandem mass spectrometry and high-performance liquid chromatography-accelerator mass spectrometry analysis have shown that the dose-response for the induction of N7-(2-hydroxyethyl)guanine) (N7-HEG) adducts in the livers of rats treated i.p. with EtO is at best described as linear, with significant increases over the background being observed at the four higher i.p. doses.

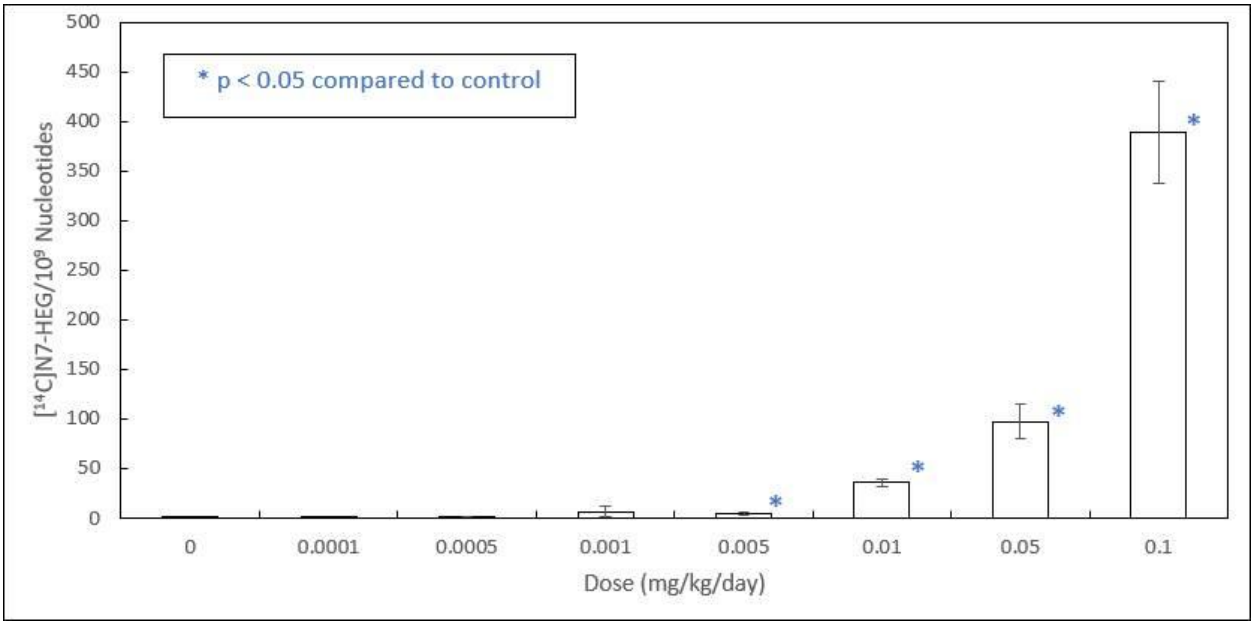

Figure 6. Dose-response for exogenously derived DNA adducts in liver of [^{14}C]EtO-treated rats measured by LC-MS/MS (Marsden et al., 2009).

Although the N7-HEG adducts are not considered mutagenic, they are the most abundant DNA adducts formed following EtO exposure (Walker et al., 1992). Thus, the shape of the dose-response curve for the N7-HEG adduct can be considered as the worst-case scenario for EtO-

induced adducts, including the most mutagenic O^6-HEG adduct whose abundance is approximately 300 times lower than that of the N7-HEG adduct (Walker et al., 1992). In reality, the slope for the mutagenic O^6-HEG adducts is expected to be much shallower than that for N7-HEG because of the kinetics of their formation and repair (Swenberg et al., 2008). Accordingly, the dose-response pattern for the molecular initiating event leading to EtO-induced mutagenicity is expected to be non-linear or at best linear at the low end of the dose-response curve and the efficiency of adduct formation increasing at higher exposure levels due to saturation of DNA-repair processes.

Further evidence for the implausibility of a steeper slope initial slope in EtO dose-response comes from genotoxicity and carcinogenicity studies conducted with ethylene. Since ethylene is metabolized in vivo to EtO, it forms the same type of protein and DNA adducts as EtO. Based upon a physiologically based toxicokinetic model, Filser and Klein (2018) predicted that exposures to 10,000 ppm ethylene induces adduct levels equivalent to EtO exposures to 3.95 (mice), 5.67 (rats), or 0.313 ppm (humans). Ethylene is not an in vivo genotoxicant in the rat or the mouse (Vergens and Pritts, 1994; Walker et al., 2000). In a chronic bioassay, ethylene was not carcinogenic in male and female Fischer 344 rats following exposed 6 hr/day, 5 days/week, for up to 24 months to concentrations of 300, 1000, or 3,000 ppm (Hamm et al., 1984). DNA adducts resulting from 300, 1000 or 3000 ppm ethylene are shown to be quantitatively to 2.4, 5.3 and 5.5 ppm EtO, respectively (Filser and Klein, 2018). Lack of ethylene carcinogenicity in the rat, in spite of increased DNA adducts equivalent to low ppm-level EtO exposure, informs that the potency of EtO's carcinogen is not higher at the lower exposures, an observation contrary to the prediction based on the IRIS (EPA, 2016a) 2-slope exposure-response model. On the other hand, the dose-response for EtO carcinogenicity is conservatively consistent with a default linear risk model with a single slope.

Fennell and Brown (2001) showed that blood concentrations of EtO in mice, rats, and humans increased linearly with exposures between 50 and 200 ppm (see figure below). Dose-disproportionate increases in blood EtO occurred only in mice at exposures exceeding 200 ppm and were attributed to substantial depletion of GSH, which limits the overall GSH conjugation capacity. It needs to be emphasized that the dose-disproportionate response in mice involved an increased (not decreased) rate of blood EtO concentration at exposures >200 ppm EtO. These observations do not support the plausibility for a steeper slope at lower exposures either for genotoxicity or carcinogenicity.

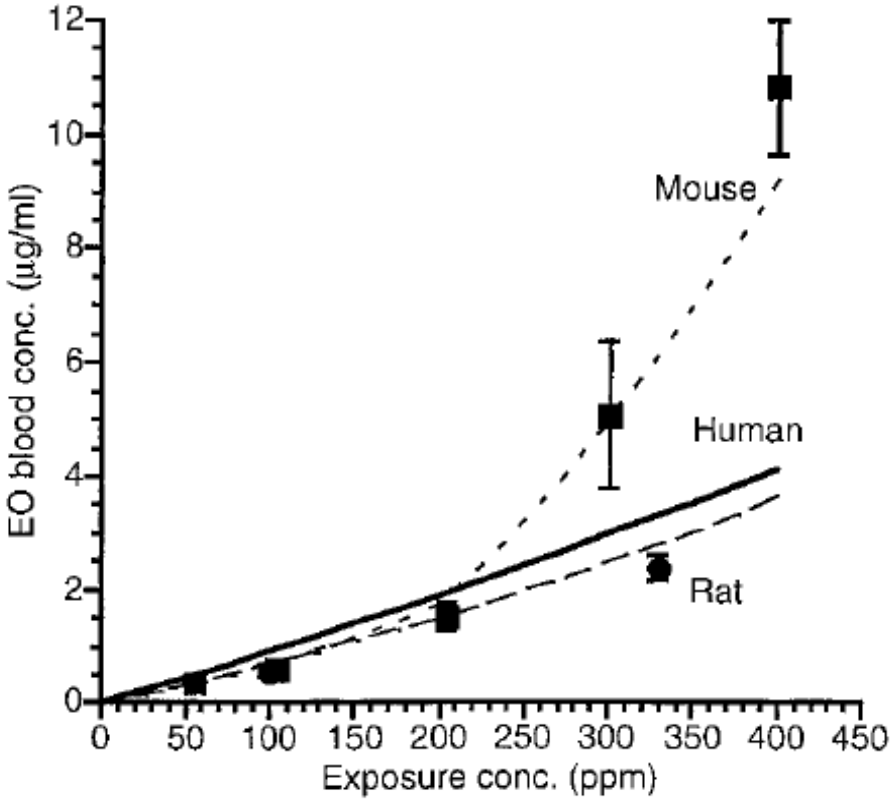

Figure 7. Toxicokinetics of EtO from Fennell and Brown (2001).

In conclusion, EtO is a relatively weak genotoxicant and requires relatively high and prolonged exposures to induce mutagenicity. The experimentally observed dose-response patterns for mutagenicity/carcinogenicity show that the CPH model is biologically more plausible than the IRIS (EPA, 2016a) 2-slope model. Accordingly, the CPH model should be the model of choice for risk assessment purposes especially if an alternate model is not a better fit to the observed data.

<u>Qualitative and quantitative analysis of genotoxicity data by Gollapudi et al. (2021) provide independent converging evidence supporting TCEQ's quantitative risk assessment.</u>

The recent literature search conducted by OEEHA missed the publication by Gollapudi et al. (2021). These authors analyzed the dose-response data to identify a point-of-departure for EtO-induced in vivo genotoxicity from an exhaustive list of published studies that employed various endpoints, tissues, and species and derived 238 ppt as the lowest permitted daily exposure (PDE) from this analysis. The PDE of 238 ppt proposed in this publication is more than three orders of magnitude higher than the 0.1 ppt established by the EPA (2016a) and similar to the 240 ppt estimated from TCEQ (2020a) UR values for 1-in-10^6 (1/M) extra risk. Thus, if EtO were operating through a genotoxic MoA for its carcinogenicity, then the exposure-response model

used by the TCEQ, rather than the one used by the EPA, is consistent with the biology based on the analysis by Gollapudi et al.

Regarding Carlsson et al. (2017) and Zeljezic et al. (2016) genotoxicity studies

The results from the studies by Carlsson et al. (2017) and Zeljezic et al. (2016) identified in the recent literature search by OEEHA should be interpreted with caution since the subjects in this study were exposed to multiple carcinogenic/genotoxic chemicals including EtO and hence it is difficult to attribute the effects observed to any single chemical. Accordingly, the statement by OEEHA that the results "……. are consistent with the overall evidence for the genotoxicity of ethylene oxide" is an overstatement.

5. OEHHA appears to equate the 5 categorical rate ratios with the 53 rate ratios, and/or assumes that comparisons of the exposure-response curve can be compared visually with the EPA IRIS (2016a) graphical representation 5 categorical rate ratios.

** Pertaining to NSRL p. 26 **

OEHHA relies on IRIS's visual fit comparisons using figures that compare continuous models with categorical models, as if the categorical model with only 5 data points were the gold standard for understanding the shape of the exposure-response curve. While categorical models with a small number of odds ratios can be useful for identifying possible associations, they do not identify the shape of the dose-response curve based on continuous data modeling as shown in detail by Valdez-Flores and Sielken (2013). Based on these visual fit comparisons, OEHHA concludes that models either over or underpredict the categorical model, which are not the data modeled.

TCEQ provided new information that OEHHA may not have been aware of. Graphical display of data is subject to manipulations including choice of how data are expressed on the y-axis and resolution of categorical models to represent the underlying individual data that were modeled. TCEQ's purpose is best expressed in TCEQ's response to peer review comments (TCEQ 2020b, p.49. 51):

> "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 MOAs and various statistical diagnostics of model fit (i.e. AIC and p-values, statistical analyses of model accuracy), consistent with the comment."

TCEQ's explains in text and illustrates in figures that EPA's graphs are misleading because EPA uses the categorical modeling results (which are not the primary data being modeled) to visually evaluate the fit of models as though these cruder categorical estimates represent the true underlying dose response. EPA correctly points out that each of these individual case categories will have very wide confidence interval (CI) but fails to address TCEQ's major point which is that the categorical estimates graphed as point estimates without the CI are not representative of the underlying 53 hazard rates modeled. EPA IRIS did not exhibit the wide confidence intervals associated with the EPA's categorical model in the graphs used to illustrate visual fit. Thus, TCEQ produced similar figures without the CI's to better illustrate the underlying individual hazard rates that are being modeled. TCEQ explains why comparison of the exposure-response model results to the categorical model results is inappropriate on p. 52 of TCEQ (2020b) response to peer review.

> "This is because while assessing model fit by visual inspection to the underlying modeled datapoints is a commonly used technique. . ., 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."

EPA counters TCEQ that the categorical model is a well-accepted method to represent the data. This true statement is irrelevant to addressing TCEQ's main point that visual fit based on categorical models are not appropriate for the purpose of determining goodness of fit of the model to the underlying data, because the categorical model is NOT modeling the underlying individual data. In addition, TCEQ cites Valdez-Flores and Sielken (2013) which is a peer-reviewed paper that demonstrates how the shape of categorical results can change with different number of categories. TCEQ is not advocating the use of these graphs to assess visual fit as a method to select the models, but instead is informing that these graphs should not be used for visual fit comparison. Instead, TCEQ relies on a more objective statistical modeling approaches to evaluate goodness of fit, rather than "eyeballing" comparisons using figures that are not fit for this purpose and distort the true comparisons of models against the underlying individual data that were modeled.

OEHHA should correct their discussion of visual fit so that all of OEHHA's claims of over or underprediction are omitted because they violate EPA IRIS' warning that such comparisons along the y-axis are incorrect comparisons of over- or under-prediction.

6. OEHHA dismisses TCEQ's prediction analysis which is a much more objective method than visual fit to check how well each of the two models (i.e., TCEQ's vs EPA's) can predict the observed number of lymphoid cancer deaths (the key cancer endpoint). In this model ground-truthing exercise, the TCEQ model better predicted not only the overall number of lymphoid cancers in the NIOSH cohort but also the observed cancers below the knot in Quintile 2.[15]

** Pertaining to NSRL p. 38-39 **

As described above in our general comments, OEHHA dismissed TCEQ's "reality check" (Table 4) based on a prediction analysis because the models are applied to general population background cancer rates, which OEHHA considered to be a flawed analysis because the NIOSH study is based on a specific cohort of occupational workers. Yet, this approach is essentially the same approach used by IRIS (EPA, 2016a, Section 4.7) to estimate extra risk for various occupational exposure levels by applying the model to the general population background cancer rates in the life-table program.

Table 4: Total NIOSH cohort lymphoid cancer mortalities predicted by TCEQ (2020a) and EPA IRIS (EPA, 2016a) models

Model (15-yr lag, MLE)	Slope Parameter (per ppm-day)	Predicted if the Model were True	100% × Ratio: Predicted / Observed	100% × SMR: Observed / Predicted	95% Poisson CI if the Model were True
TCEQ (CPH)	2.81E-06	52.42	98.9%	(40.1, 70.0)	(38.2, 66.6)
IRIS 2-slope spline 15-yr lag (MLE)	**7.58E-04**	**91.69**	**173.0%**	**(70.1, 122.4)**	**(72.9, 110.4)**

Note: There are 53 actual lymphoid mortalities. 53 is within the CIs for the TCEQ model but not within the CIs for the IRIS model. Thus, the TCEQ model accurately predicts the actual cancers. In contrast, the IRIS model statistically significantly (bold font) over-predicts the actual number of cancers. TCEQ used the inverse of the confidence intervals of the SMRs. We calculated the confidence intervals based on the Poisson distributions. See TCEQ (2020a, Table 6).

[15] TCEQ (2020a) defined Quintile 1 as the 9 lagged-out cases (no exposures). The remaining 44 cases were equally divided into 4 groups designated by TCEQ as Quintiles 2-5. EPA IRIS (2016a, p. 4-15) reported 13 exposed cases below the knot of 1600 ppm-days. Thus, prediction of Quintile 2 comprised of 11 cases with the lowest exposures best reflects "local" fit below the knot.

EPA (2022) also raised a question about TCEQ's calculation of the confidence interval (CI) based on the inverse of the SMR CI as described in detail by TCEQ (2020a). Thus, it is useful to show that another well-accepted approach for estimating the 95% CI results in the exact same conclusion that the TCEQ model accurately predicts the actual 53 cases, whilst the IRIS model over-predicts the number of cases (Table 4)

OEHHA claims that TCEQ calculations did not accurately account for any differences that might exist between the general US population and the NIOSH cohort. It is unclear if OEHHA is aware that the TCEQ (2020a) DSD[16] includes a sensitivity analysis to demonstrate that the TCEQ model better predicts the overall actual cancers even after applying a high HWE of 15-16% for lymphoid cancers as a sensitivity analysis (Table 5).

Table 5: Total NIOSH cohort lymphoid cancer mortalities predicted by TCEQ (2020a) and EPA IRIS (EPA, 2016a) models with 15% HWE as a sensitivity analysis

Model (15-yr lag, MLE)	Slope Parameter (per ppm-day)	Predicted if the Model were True	100% × Ratio: Predicted / Observed	95% Poisson CI if the Model were True
TCEQ (CPH)	2.81E-06	44.56	84.1%	(31.4, 57.6)
IRIS 2-slope spline 15-yr lag (MLE)	**7.58E-04**	**77.94**	**147.1%**	**(60.6, 95.2)**

Note: The TCEQ model still accurately predicts the actual cancers after accounting for a theoretical HWE. In contrast, the IRIS model statistically significantly (bold font) over-predicts the actual number of cancers after including a theoretical HWE[17]. TCEQ used the inverse of the confidence intervals of the SMRs. We calculated the confidence intervals based on the Poisson distributions.

Although one can quibble with the TCEQ's selection of 15-16% based on a Norwegian worker study with relatively short average follow-up of 11.5 yrs (Kirkeleit et al. 2013),[18] the larger point is that a 15% HWE is a very reasonable high estimate for any differences that might exist between the general US population and the NIOSH worker. The NIOSH study authors themselves concluded that there was unlikely to be a cancer HWE in this longer follow-up study (Steenland et al., 2004[19]) cohort. This conclusion of the NIOSH study authors is very consistent with the general experience in cancer epidemiology that HWE is known to vary with type of

[16] TCEQ Section A3.3.2

[17] Predicted is based on multiplying predicted values in Table 4 by 0.85 for HWE of 15%, and CI's calculated using Poisson distribution. Compare with TCEQ (2020a, p. 102, Section A3.3.2) estimates of 44.3 (95% CI: 33.9, 59.2) and 77.5 (95% CI: 59.3, 103.6) based on 15 and 16% HWE for males and females, respectively.

[18] Kirkeleit et al (2013) did not find a HWE for lymphoid or breast cancer. It is unknown if the Norwegian cohort is representative of the NIOSH sterilizer workers.

[19] "The healthy worker effect would seem an unlikely explanation for the lack of cancer excesses in the exposed versus non-exposed comparisons." (Steenland et al., 2004)

disease, being smaller for cancer than for other major diseases, and it tends to disappear with time since recruitment into the workforce (IARC, 1999). In addition, the epidemiologic literature has shown that a HWE is predominantly related to populations with shorter follow-up and non-cancer causes (Monson, 1986; Fox and Collier, 1976).

Using a 15% HWE, the CPH model accurately estimates the observed number (53) of lymphoid deaths in the NIOSH study (Table 5). In contrast, the linear 2-piece spline model statistically significantly overestimates the number of observed lymphoid deaths in the NIOSH study (Table 5). In addition, a quintile analysis was also performed by TCEQ (2020a) to address EPA IRIS (2016a) emphasis on the local fit of the models below the knot. EPA IRIS (2016a, p. 4-15) reported 13 exposed cases below the knot of 1600 ppm-days. Thus, prediction of Quintile 2 comprised of 11 cases best reflects "local" fit below the knot. Table 6 summarizes TCEQ CPH and EPA 2-piece spline model predictions of the number of lymphoid deaths at each quintile. Table 6 shows that, for each quintile, the CPH model has superior local fit These results indicate that the CPH model not only has better local fit below the knot, but also at the highest quintile.

Table 6: Quintile-specific NIOSH cohort lymphoid cancer mortalities predicted by Cox and linear two-piece spline models

Model	Quintile 2	Quintile 3	Quintile 4	Quintile 5
Lymphoid Cancer Deaths Observed in NIOSH Cohort	*11*	*11*	*11*	*11*
Standard Cox model – 15-yr lag (MLE)	14.4 (8.1, 28.9)[20] (7, 21.8)[21]	8.0 (4.5, 16.1) (2.4, 13.5)	9.4 (5.2, 18.8) (3.3, 15.4)	9.1 (5.1, 18.3) (3.2, 15.0)
Linear two-piece spline with knot @ 1,600 ppm-days – 15-yr lag (MLE)	**20.9** **(11.7, 42.0)[20]** **(11.9,29.8)[21]**	17.6 (9.8, 35.2) (9.3, 25.8)	**20.8** **(11.6, 41.7)** **(11.8, 29.7)**	**20.9** **(11.7, 41.9)** **(11.9, 29.8)**

Note: The TCEQ model accurately predicts the actual cancers for the lowest exposure quintile 2. In contrast, the IRIS model statistically significantly (bold font) over-predicts the actual number of cancers. TCEQ used the inverse of the confidence intervals of the SMRs. We calculated the confidence intervals based on the Poisson distributions. See TCEQ (2020a, Table 6).

In conclusion, the TCEQ standard CPH model accurately predicts the number of lymphoid deaths observed in the NIOSH study while EPA's two-piece linear spline model statistically

[20] TCEQ (2020a) method was used for CI, for comparison we calculated the CI based on Poisson distribution.
[21] CI calculation based on normal distribution for all Quintiles.

significantly (at the 2.5% significance level) overpredicts the number of lymphoid deaths observed in the NIOSH study. This is true with and without consideration of a reasonably high HWE of 15% which reasonably accounts for any differences that might exist between the general US population and the NIOSH worker cohort given the absence of a HWE in the NIOSH cohort. This TCEQ "reality check" is a well-accepted approach that is essentially the same approach used by IRIS (EPA, 2016a, Section 4.7) to estimate extra risk for various occupational exposure levels by applying the model to the general population background cancer rates.

7. **OEHHA's sole reliance on internal analyses and OEHHA's extreme and complete exclusion of external analysis is based on a main conclusion that all external analysis should be ignored because it is confounded by the HWE. This is contradicted by NIOSH study authors' own published conclusions that "the healthy worker effect would seem an unlikely explanation for the lack of cancer excesses in the exposed versus non-exposed comparisons." Furthermore, OEHHA's uncritical acceptance of conclusions of a more recent paper by Park (2020) that there is a Healthy Worker Survival Effect (HWSE) led OEHHA to support the EPA IRIS model. However, the conclusions are not supported by the actual results in the paper. Improvements in these and other descriptions of the human epidemiological studies are needed to accurately assess the epidemiological weight-of-evidence.**

** Pertaining to NSRL p. 30 **

External (comparisons to the general population) and internal analyses (worker to worker comparisons) in occupational epidemiology studies are complementary approaches to the examination of potential exposure-response associations. When they agree, confidence in the presence or absence of risk is enhanced. When they disagree, it is incumbent upon the researchers to explore explanations. In some cases, and more often in the early years of occupational epidemiology, the identification of the potential for the HWE led to a general preference in favor of internal analyses. The HWE, particularly for cancer outcomes, is of much lesser concern if the external analyses of cohort studies have been updated with longer observation period. We have commented on this being the case for the NIOSH and UCC studies. The UCC study has had two published updates since the original Greenberg et al. (1990) publication (Teta et al. 1993, Swaen et al. 2009), such that the average follow-up of study subjects has gone from 20 to 37 years. This is what Steenland et al. 2004 was referring to below when he noted the change observed from the original study published in 1991.

> "The healthy worker effect has diminished (all-cause mortality was up to an 0.90 from the prior SMR of 0.81) as would be expected with increased follow up."

ACC's analysis and explanation of the HWE is also supported by IARC (1999) in its textbook on Cancer Epidemiology: Principles and Methods, which specifically notes that HWE "is known to

vary with type of disease, being smaller for cancer than for other major diseases, and it tends to disappear with time since recruitment into the workforce."

Furthermore, comparisons to worker populations can have their own limitations, such as small sample sizes, baseline risks that suggest they are non-representative of the true low or non-exposed population (e.g., Mikoczy et al., 2011).

OEHHA's complete reliance on internal analyses leads them to consider well-conducted meta-analyses as "flawed" and ignore their conclusions. In general, meta-analysis is a well-accepted method for summarization of results from multiple studies, despite being generally limited to pooled overall risk estimates, as many studies do not provide data by levels of exposure. Furthermore, OEHHA is incorrect in saying that Marsh et al. (2019) did not consider the results of NIOSH internal analyses. In fact, these authors, state:

> "However, similar to the LHC results, the NIOSH findings, which revealed no overall excess for breast cancer, were limited to the exposure–response analyses using the log-transformed EtO exposure metric and were questioned by the authors due to their inconsistency across the other EtO metrics considered and potential case over-ascertainment in the higher exposure categories. As discussed above, due to the questionable validity of the positive EtO exposure–response of Mikoczy et al. (2011), these fndings add little weight to the overall evidence for EtO exposure and breast cancer."

It is, therefore, scientifically unjustified to ignore the absence of any overall excess of breast cancer in the NIOSH 2004 mortality study and the relevant meta-analyses (Marsh et al. 2019; Vincent et al. 2019).

OEHHA's crude and incomplete evaluations and uncritical acceptance of Park, 2020 led them to unequivocally support the EPA model.

** Pertaining to NSRL p. 33-34**

OEHHA states that they conducted several qualitative and quantitative assessments of potential bias and errors in the NIOSH study and its use by EPA.

1. Their discussion of exposure assessment is limited to the successful validation NIOSH performed of the post 1978 estimates from their regression model. OEHHA does not consider the NIOSH limitations related to exposure estimates for the pre-1978 period that had no validation, as ACC has previously discussed in detail. Instead, they criticize Bogen et al. (2019) for not validating their estimates.

2. While OEHHA raises the issues of the healthy worker effect (HWE) and the healthy worker survivor effect (HWSE), they conclude that the internal analyses by NIOSH remove the HWE issue and there is no more than a minor downward OR bias (10% or less), due to the HWSE. They cite Park 2020 to support the existence of a HWSE in the NIOSH data. We agree with OEHHA that neither of these types of biases are relevant. However, there is no HWE in the NIOSH study due to extensive follow up, as noted by Steenland, which adds to the relevance of the external analyses of these data. The absence of findings in external analyses of the NIOSH (and UCC data) adds to the uncertainty of the EPA model suggesting a highly potent carcinogen.

3. We also disagree with citing Park 2020 in support of the existence of a HWSE in the NIOSH study, such that control for employment duration leads to a stronger association between cumulative exposure and breast and hematopoietic cancers. In Section 3.2 of Park 2020, the author discusses the findings reported in Table 2. The models include both cumulative exposure to EtO and duration of employment – variables that are very likely highly correlated. Park reports "statistically significant negative effect of duration (diminishing rate of leaving with increasing time on job) and positive effects of EtO cumulative exposure that are highly significant for all but the smaller work group of black women (Table 2)." These two variables are in the model used to fit the data with a similar multiplicative role. Given the high correlation between cumulative exposure and duration of employment, a negative coefficient for duration of employment would have to be compensated by a positive coefficient for cumulative exposure. Thus, it is not surprising that a negative coefficient for one variable results in a negative coefficient for the other variable. In fact, in three of the analyses reported in Table 2, the more negative of the coefficient for duration resulted in more positive coefficients for cumulative exposure.

   Park (2020) Table 5 shows the results of Park's models for female breast cancer. There, Park does not find any statistically significant relationship between breast cancer and EtO exposures or employment duration. Park states "For the 102 deaths from female breast cancer, there was no statistically significant difference in mortality on cumulative EtO exposure with a 10-year lag." Park goes on to indicate that "with 20-year lag, the contribution of cumulative EtO was significant." It is interesting that although the model is significant, the 95% confidence intervals indicate that the coefficients for cumulative exposures to EtO are not statistically significantly different from zero. That is, the 95% confidence intervals indicate that breast cancer in female workers is not related to cumulative exposure to EtO lagged 20 or 10 years.

   Park (2020) Table 6 shows the results for lymphopoietic cancer deaths in the male and female workers. There, Park showed the results for male black workers only because "The 73 lymphopoietic cancer deaths did not represent an overall excess (SMR = 0.96; 95% CI, 0.76-1.20), based on U.S. rates, particularly for white men (SMR = 0.92), and white women

(SMR = 0.85), but among black workers, there was a statistically significant increase in SMR with cumulative EtO exposure (lagged 2.5 years; LRT: P = .011) (Table 6; model 1)." Similar to the results in Table 5 for breast cancer, in Table 6 the models statistically significantly improved the model fit to the lymphohematopoietic deaths in black workers but the coefficients for the cumulative exposure to EtO were not significantly different from zero; that is, there is no statistically significant increasing relationship between lymphopoietic cancer and cumulative exposure to EtO lagged 10 years even in the most sensitive subgroup of workers in the NIOSH study.

The above data are inconsistent with Park's description of his results in his abstract that were uncritically accepted by OEHHA.

4. OEHHA attempted to address the possibility that high intensity- short exposures might explain the NIOSH findings and therefore not be generalizable to the general population, not subject to such types of exposures. Without actual NIOSH data and with incorporation of several questionable assumptions, OEHHA concludes that excluding workers with these exposures from the NIOSH study would have little impact on the EPA exposure-response slope. Firstly, they assumed that workers with this type of exposure would most likely be in the middle categories of cumulative exposure. Then OEHHA estimated case and control counts in each exposure category, recalculated ORs and exposure-response slopes after excluding various percentages of participants (e.g., 10–30% high intensity-short duration exposed workers) in the middle exposure categories, using guestimates since the actual data were not available. "Exclusions were done at the case:control ratio equal to or slightly lower than that reported in the highest exposure category (where almost all workers probably had at least some high intensity exposure). Overall, these exclusions (with and without replacing the excluded participants into the highest category) had little impact on exposure-response slopes (e.g., 10% or less). This suggests that this issue did not have a major effect on the unit risk calculations or the generalizability of the NIOSH findings." This conclusion is hardly justified using this obscure analysis and exclusion assumptions, in the absence of actual NIOSH data.

We agree with OEHHA's summary of the limitations of community-based studies, which makes them inappropriate for exposure-response but disagree that they are useful for hazard assessment.

** Pertaining to NSRL p. 9, 37 **

The proposed NSRL (p. 9) states:

In addition to the three human epidemiological studies in Table 1, OEHHA identified four epidemiological studies that investigated associations between residential proximity to

ethylene oxide emitting facilities and increased cancer risk (Garcia et al., 2015; Bulka et al., 2016; Hart et al., 2018; and Jones et al., 2023). Emissions data were obtained at the community level from US EPA databases: the Toxics Release Inventory (TRI) (Bulka et al., 2016; Jones et al., 2023; US EPA, 2023a) and the National Air Toxics Assessment (NATA) Garcia et al., 2015; Hart et al., 2018; US EPA, 2018 . While these community-based air pollutant studies can be useful for hazard identification, OEHHA judged them to be less useful for dose-response assessment of ethylene oxide compared to the occupational studies (Steenland et al., 2003 and 2004; Swaen et al., 2009; and Mikoczy et al., 2011) due to greater uncertainty in estimating individual exposures. This can result in non-differential exposure misclassification and bias risk estimates towards the null (Shy et al., 1978). Furthermore, there were fewer exposed cases, and there may be less exposure contrast in these community-based studies of ethylene oxide, decreasing the sensitivity of the studies to detect an effect."

We agree with the limitations noted by EPA but disagree that community-based studies are useful for hazard identification of EtO, precisely because of limitations noted in EPAs final sentence. But in addition to concerns about sensitivity (study power), there are serious concerns in such studies related to confounding risks that may be related to potential errors in exposure assessment due to confounding by other sources of EtO exposure, such as smoking and highways. Bias can, therefore, be in either the positive (specificity) or negative direction (sensitivity), making such designs uninformative.

OEHHA fails to discuss the most important limitation of the Mikoczy et al., 2011 study – the questionable comparison group in the internal analyses

** Pertaining to NSRL p. 12 **

OEHHA has but one criticism of Mikoczy et al (2011): "Exposures were much lower than in the NIOSH and Union Carbide cohorts, which decreases the ability to detect an effect." This is apparently OEHHA's explanation for the failure in this study to detect any increase in LH cancers.

While not noted by OEHHA, Mikoczy et al. (2011) has been incorrectly cited in IRIS as supportive of a supralinear association with breast cancer, despite an overall deficit of breast cancer (SIR= 0.81), with or without consideration of a latency period. However, the two higher cumulative exposure groups had statistically significant elevated rates of breast cancer, due to a substantial and statistically significant deficit of breast cancer in the low dose reference group. This deficit is not explained by the HWE, which is primarily related to non-cancer causes and declines with length of follow up. As discussed above, there are clearly advantages to comparing workers to workers in epidemiology studies to overcome possible biases in external

comparisons to the general population. However, there may also be disadvantages to using an internal comparison group that may not be recognized. One danger is selecting a referent group that has an unusual deficit of the disease of interest that creates an artifact of an excess as is illustrated in this study, whose referent group breast cancer rates are 50% of general population baseline. This serious limitation was also illuminated in Marsh et al. 2019:

> "The validity of the Mikoczy et al. (2011) finding and conclusion can be challenged, however, on the basis of several methodological issues. First, the greater than 2-fold relative excesses in breast cancer incidence risk in the two highest cumulative EO exposure categories were ensured by an inordinately large, statistically significant 48% deficit in breast cancer incidence in the baseline category. The inordinately low baseline SIR for breast cancer is puzzling given that regional rates were used in the external comparisons and that there was no apparent problem with under-ascertainment of breast cancer cases. The healthy worker effect is also not a reasonable explanation for the low baseline breast cancer rate (Gridley et al. 1999). It appears that for unknown reasons, the baseline group used by Mikoczy et al. (2011) differs from the highest two cumulative EO exposure groups on factors other than EO exposure that may be related to breast cancer."

The EPA (2016a) IRIS report quantitatively demonstrated the inconsistency of the excesses reported at very low exposures in this population with excesses at only higher exposures in the NIOSH study.

> "Thus, crude comparison analyses were done to evaluate whether or not the exposure-response models of the NIOSH study that were used to derive unit risk estimates in this assessment gave predictions consistent with the Mikoczy et al. (2011) internal incidence ratios (IIRs) for the two highest exposure quartiles (see Section J.2.2 of Appendix J). The predicted values for lymphoid cancer were within the 95% CIs for the IIRs for lymphohematopoietic cancer reported by Mikoczy et al. (2011). The predicted values for breast cancer incidence, however, were below the lower limit of the 95% CIs for the IIRs for breast cancer, suggesting that the Mikoczy et al. (2011) results are consistent with a higher unit risk estimate for breast cancer incidence than the one derived in this assessment. The reasons for the discrepancies are unknown… "(EPA 2016a, p.60-61)

Marsh et al. 2019 also addressed this exposure inconsistency:

> "Second, cumulative EO exposure levels in the Mikoczy et al. (2011) study were very low relative to both the UCC cohort (Swaen et al. 2009) and NIOSH breast cancer cohort incidence study (Steenland et al. 2003)."

EPA more recently argues that there were some substantial exposures in the Mikoczy et al. cohort, making it more comparable to the exposures in the NIOSH cohort, backtracking from the IRIS report. OEHHA reports the breast cancer findings from the internal analysis and ignores the potential bias associated with a non-representative worker comparison group.

8. **Based on uncritical acceptance of the IRIS evaluation, OEHHA inaccurately exaggerates the reliability of the NIOSH worker exposure estimates prior to 1978.**

** Pertaining to NSRL p. 23, 24, 34 **

OEHHA can improve the weight of evidence of the epidemiological data by more accurately describing the substantial limitations of the NIOSH worker exposure estimates prior to 1978. Furthermore, NIOSH has lost the electronic files needed to independently assess the NIOSH estimates of exposure prior to 1978 for which NIOSH had no direct measurements[22]. These data quality issues make open access and independent assessment of the IRIS cancer slope factor difficult for lymphoid cancers and impossible for breast cancers.

OEHHA makes the following statement regarding the NIOSH cohort exposure assessment,

> "*Each participant's EtO exposure was estimated using a validated multiple regression exposure model that incorporated information on workplace air measurements, sterilization unit size, engineering controls, timing of sterilization, product type, calendar year, and historical process changes."* (Proposed NSRL p. 23)

OEHHA also states,

> *"US EPA judged the NIOSH study to be of "high quality" based on the availability of quantitative exposure estimates for individual workers, high-quality exposure assessment, longitudinal study design, large sample size, inclusion of males and females, adequate follow-up, absence of known confounding exposures, multiple study locations, and the use of internal comparison groups. OEHHA reviewed the NIOSH study using the Hill guidelines for causal inference and the National Toxicology Program (NTP)'s risk of bias tool, and also concluded that this study is of high quality, and unlikely to be affected by important bias or confounding."* (Proposed NSRL p. 24)

OEHHA dismisses a robust analysis of the trend of the NIOSH exposure data by Bogen et al. (2019):

> *"Bogen et al. (2019) have suggested that exposures occurring prior to 1978, the first year that EtO sampling data were available for the NIOSH cohort, may have been dramatically under-predicted by the NIOSH exposure model. However, as noted by these authors, several assumptions were used in their assessment, and the information used to support these assumptions, "were limited in scope and quantitative detail." In addition,*

[22] In response to the panel's suggestion that the Hornung analysis represents an "invaluable opportunity" for further analysis of the impact of possible errors in exposure estimation, the EPA investigated the possible use of the "errors in variables" approach (page 27 of the panel report). . . Unfortunately, the electronic data files used in the exposure analysis were no longer available, so that analysis based on the errors-in-variables approach was not possible.

the authors were unable to validate their pre-1978 predictions since no actual worker measurements were available from that time. Overall, because of these and other weaknesses, the accuracy of the Bogen et al. (2019) assessment is unknown." (Proposed NSRL p. 34)

These OEHHA statements show, at a minimum, a lack of rigor in evaluating the NIOSH exposure model (Horning et al. 1994) and its "validation", and a biased evaluation of the Bogen et al. (2019) paper as discussed in our general comments above.

A more correct statement describing the NIOSH multiple regression exposure model is that the model was only validated for the period 1979-1985 (very few samples were collected from 1976-1978) during which EtO concentration measurements were collected from six facilities, but not for the earlier period of sterilizer facility operations (Pre 1978; from late1930's to late 1970's) when a majority of cohort workers were occupationally exposed to EtO and no exposure measurements were collected. An incorrect assumption by Hornung et al. in applying the NIOSH model to the Pre1978 sterilizer operations regarding a key variable "calendar year", a surrogate for improvement in work practices, inferred no changes in the Pre1978 period and raises serious question about reliability of the NIOSH model in prediction early worker exposures and the dose-response relationship based on the exposures of these early workers in this cohort.

There are several reasons to question the reliability of NIOSH model predictions of early sterilizer worker exposure estimates:

- First, the NIOSH multiple regression exposure model validation (Hornung et al. 1994) was based using a portion of the 1978-1986 to develop the model and another portion of the data to test the model. Therefore, the model was validated for the years with data but not for the early years for which no concentration measurements were available. Hornung et al. provided no other data, information or analyses for sterilizer workers or processes in this earlier time period to check the applicability to this NIOSH exposure model to sterilization workers in the cohort who worked from the late 1930's to later 1970's. During this time period the vast majority of cohort worker and fraction with most reported cancers. These facts regarding the NIOSH model validation should have raised concern about model predictions for early sterilization workers.

- Second, Hornung et al. selected Calendar year, a surrogate for improvement in work practices, as the key variable in the NIOSH exposure model to Pre1978 cohort workers. This variable was applied conditionally on the max year 1978 forcing the inverse parabolic fit for years before 1978. This application of calendar year inferred that there were no major changes in work practices in sterilization operations between the late 1930's and late 1970's that would have affected worker exposures. Contrary to this

inference, substantial published information and data on early work practices and changes in work practices were found in technical literature, industry documents, and from early workers and industry experts (Bogen et al. 2019). In other words, Bogen et al. (2019) brought far more information and data to inform the validity of the NIOSH model prior to 1978, when NIOSH had insufficient (1976-1978) or no (<1976) data. For example, the numbers of repeated cycles of in chamber, post-exposure vacuum air- or nitrogen washes have increased from two or fewer from early operation up to ten or more for operations in the 1980's leaving high levels of EtO residues to off gas in from sterilized materials and packaging from early operations and lower levels in later operations (Goldgraben and Zank 1981; Buonicore et al. 1984). Consistent with few wash cycles, there are several published studies of rates of EtO off gassing from sterilized materials representing conditions in the 1950's through 1980's (Bruch 1961, 1972; Buonicore et al. 1984; FDA 1978; Stetson et al. 1976; White 1977). As importantly, early operation stored sterilized materials in the same room as ongoing sterilizer operations where both operational emissions and sterilized material off gassing contributed to worker exposure while later operations moved sterilized material to a separate warehouse room reducing the exposure of highly exposed sterilizer operators (Bogen et al. 2019). Clearly there were important work practice changes over time that need be considered in assessing the exposure of cohort workers.

- Third, the NIOSH exposure model based on the conditioning of calendar year predicted early sterilization workers were exposed to EtO concentration substantially lower than workers in 1978 when exposure concentration predictions were based on measurements (see Figure 8 below from Bogen et al. 2019). One would not expect low exposure concentration when equipment and process were crude, and little was known about EtO toxicity and no worker protection regulations. The NIOSH model predicted early worker increasing exposure pattern is inconsistent with industrial hygiene data collected in other industries (e.g., on Grote et al. 2003, 2006) and inconsistent with historic worker exposure guidance (ACGIH 1948, 1957) for EtO exposure concentrations in the workplace. ACGIH provided an exposure limit for EtO of 100 ppm in 1948 and 50 ppm in 1957 to encourage reductions in workplace exposure. As a reality check, no ACGIH guidance would have been needed had EtO concentration been as low as predicted by the NIOSH exposure model.

- Fourth, Bogen et al. (2019) performed an engineering/industrial hygiene evaluation of early sterilization worker EtO exposure to assess the reliability of NIOSH exposure model predictions of an increasing exposure trend. Bogen et al. concluded that from the late 1930's to 1978 there was a decreasing exposure trend for sterilizer workers rather than increasing trend from very low exposures to high exposures predicted by the NIOSH exposure model (Figure 8 from Bogen et al. 2019, Figure 5).

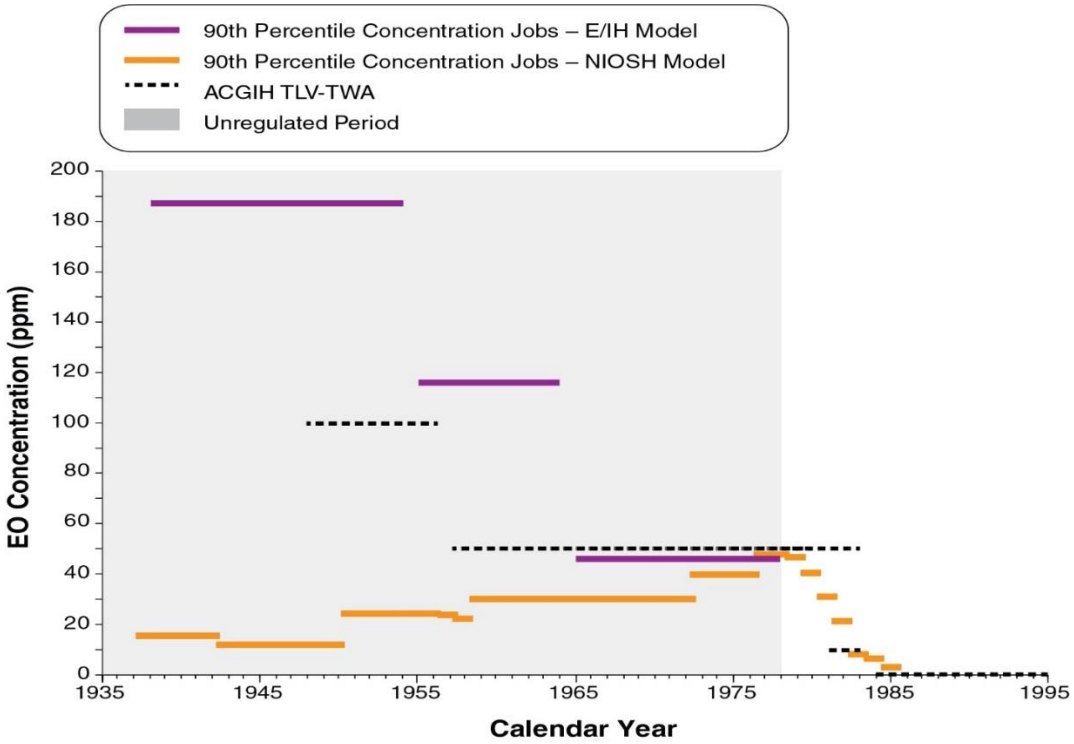

Figure 8. Comparison of E/IH (purple lines) and NSR (orange lines) exposure model estimates of occupational respiratory exposures to EtO in facilities that sterilized medical/health products and prevailing ACGIH TLV limits for EtO (dashed lines). Shaded area represents the period during which very limited or (pre-1976) no contemporaneous measurements were available to validate NSR model predictions and during which no EtO-specific regulations were in place to limit occupational EtO exposures. Adapted from Figure 5 of Bogen et al. (2019)

There are no data or analyses available to support OEHHA's agreement with EPA that the NIOSH model produced a "high-quality" exposure assessment. To the contrary, Bogen et al. (2019) published substantial data and analyses showing that the NIOSH model is flawed and there was a decreasing rather than increasing EtO exposure trend for pre1978 sterilizer operators contributing uncertainty to the EPA and OEHHA EtO risk assessments. OEHHA should recognize the limitations of the NIOSH multiple regression exposure model predictions for early sterilizer operators and the potential adverse effects it has on estimation of risk. Assigning cases with underestimated exposures means that the lymphoid cancer is associated with lower EtO levels than the workers with lymphoid cancers had been exposed to. In general, underestimating exposures associated with cancers will lead to an overestimation of potency. In addition, as there are data on worker exposures to EtO pre-1978 for production workers, it is suggested that OEHHA review production worker cohort monitoring data to ascertain a more reliable picture of early worker exposure patterns.

9. The Union Carbide Corporation (UCC) cohort should play a prominent role in considering the strength and consistency of the epidemiology data in supporting the IRIS vs. TCEQ UR. OEHHA's description of this cohort incorrectly omits the internal analysis by Valdez-Flores et al. (2010) which included exploration of different exposure metrics and lag times. The absence of findings in the UCC cohort for male lymphoid cancer is not consistent with a steep slope at low concentrations.

** Pertaining to NSRL p. 9, 11 **

This UCC study is very important to consider in the weight of evidence because it included long-term follow-up of workers from the 1940s, the infancy in EtO production. The absence of findings in the UCC cohort for male LH is not consistent with a steep slope at low concentrations and does not support the IRIS (EPA, 2016a) derivation of one of the highest IURs. This cohort of 2,174 workers was a subset of another NIOSH study of or more than 29,000 UCC chemical workers in the Kanawha Valley (KV) of WV (Rimsky et al., 1988).

Compared to the NIOSH sterilizer worker mortality study, the UCC study has a smaller cohort sample size but has comparable numbers of LH and lymphoid cancers as those reported in the male component of the NIOSH cohort. The UCC study is, therefore, informative, with respect to males, and the increases in LH and lymphoid cancers reported in the NIOSH study were limited to males (Steenland et al., 2004).

Furthermore, the quality of the UCC study exposure assessment is comparable to that of the NIOSH study. It employed individual exposure estimates for a substantial period between 1925-1988, utilizing the Greenberg et al. (1990) validated categorization of EtO producing and using departments by level of exposure and quantitative estimates of average intensity by these categories developed by Teta et al. (1993). There were no potential confounders to other chemical exposures because such workers were removed from analysis by Teta et al. (1993), Swaen et al. (and Valdez-Flores et al. (2010). Exposure data were available for study subjects at the West Virginia (WV) locations starting in 1974. They were available from Union Carbide's Texas City plant that operated identically to the WV location from the early 1960s. This represents an important advantage of the UCC study over the NIOSH study, which had no exposure data prior to 1978 (Bogen et al., 2019; see previous section for detailed discussion). Estimates from EtO operations in the literature were used for the 1940-1956 exposure period, although only a small percentage of the cohort were employed during that period.

Contrary to Table 7 of OEHHA's Proposed NSRL, there was exploration of log cumulative exposures and multiple lag times in the UCC study. Valdez Flores et al. (2010) reports that Table S11 in Supplemental materials indicates that the fit (maximum likelihood) varies depending upon the exposure scale used in the log cumulative exposure model (i.e., ppm-days,

ppm-years, ppb-days, and ppb-years). Table S11 also illustrates that the Cox proportional hazard model with the slope parameter multiplying cumulative EO exposure fits the data better than any of these four alternative log cumulative exposure models in more than 55% of the combinations of 12 endpoints. None of the other lag periods resulted in a change in statistical significance, therefore no lag was included in the Valdez-Flores et al. (2010) publication. Nevertheless, to be consistent with NIOSH, a lag period of 15 years was applied to the lymphoid analysis by TCEQ (2020a).

In summary, the absence of increases in LH and lymphoid cancers in the UCC study (in both external and/or internal comparisons), as well as the statistically significant increases in the NIOSH study limited to male highest exposure groups in internal comparisons conducted by Steenland et al. 2004, call into question the biological plausibility of the very high IRIS IUR for LH cancers. Overall, the epidemiological evidence does not support EtO as a potent carcinogen with a steep exposure-response pattern at low exposures. The standard CPH model used to derive the TCEQ IUR is a model well-accepted by epidemiologists in cancer exposure-response analysis, is linear at exposure levels of interest, and consistent with an assumption of no-threshold that reflects the epidemiological weight-of-evidence.

10. We applaud the proposed NSRL for including a section on endogenous exposure to EtO and have recommendations for improvement. While EPA's potency estimate technically only applies to exposures above endogenous levels, it is implausible that a chemical would be a potent carcinogen at levels that the body produces through natural processes. The key assumptions in extrapolating the dose-response relationships at lower exposures are scientifically valid and are now corroborated by independent data sets on smoking.

** Pertaining to NSRL p. 18, 19 **

We applaud the OEHHA NSRL and IUR supporting technical documents for including a section on endogenous exposure to EtO in their documents. This topic area is important to future risk assessment and risk management decisions made for this unique chemical. We recommend making the following improvements to this section:

The biochemical pathways that contribute to endogenous exposures, include: (1) production of ethylene by bacteria normally present in the gastrointestinal tract, which is then absorbed into the body; and (2) systemic production of ethylene by specific precursors and by oxidative stress. Endogenous production of EtO results from the oxidation of ethylene resulting from both sources. These pathways are operable in all mammalian species, with measured EtO biomarker levels (2-hydroxyethyl valine or HEV) generally being higher in laboratory rats and mice than in humans.

Endogenous exposures to EtO are variable. These exposures vary from person to person (interindividual variation) and from day to day (temporal variation), and can be modulated by diet (e.g., fatty acid composition; diet content of precursors that are metabolized to ethylene), medications (e.g., antibiotics), and underlying conditions (e.g., oxidative stress).

Kirman et al. (2021) estimated endogenously produced EtO is the largest contributor to EtO-biomarker levels in general population nonsmokers. The estimated average HEV burden of 29.2 pmol/g Hb resulting from endogenous exposure corresponds to an equivalent median inhalation exposure to 2.3 ppb EtO in air. In this context, EtO RSC exposure (0.0016 ppb) is more than 1000-fold lower than the endogenous exposure and would generally be considered negligible.

The Proposed NSRL can clarify that background exogenous exposure is generally a small fraction of total background exposure

** Pertaining to NSRL p. 18 **

As most of general population background exposure arises from endogenous production (~95%), whereas exogenous exposure via inhalation of EtO in ambient air generally constitutes a small fraction (~5%) of total exposure (Kirman et al. 2021. EtO in ambient air has been sampled since 2018 at background monitoring locations across the U.S. under the EPA NATTS and UAT hazardous substances monitoring programs. Samples also have been collected at local/regional background locations as part of monitoring programs for EtO emitting facilities. Therefore, there are substantial data characterizing general population background exogenous exposure.

The Proposed NSRL can add that total background exposure from endogenous and exogenous pathways has been characterized from CDC NHANES biomarker data and Kirman et al. equivalent concentrations

** Pertaining to NSRL p 18-19 **

Our knowledge of EtO background exposure is informed by CDC internal dose data in the form of a representative exposure biomarker, N-(2-hydroxyethyl)-valine (HEV) adduct levels, measured in erythrocytes for nonsmokers and smokers in the U.S. population (CDC 2019; Kirman et al. 2021). HEV adduct levels represent an individual's total background EtO exposure from endogenous and exogenous sources. Kirman et al. 2021 developed a relationship between biomarker (HEV) concentration and total and endogenous equivalent concentrations (equivalent continuous exposure concentrations in ppb) for smokers and nonsmokers in the U.S. population. Endogenous and total equivalent levels reflect air concentration of EtO that are equivalent to the levels that are produced endogenously, and endogenously and exogenously, respectively. Filser and Klein (2018; Figure 12A) study provides an independent PBPK model-based validation of the linear equivalent relationship adopted by Kirman et al. (2021).

Although OEHHA's Proposed NSRL (p. 19) is correct that cancer risks account for endogenous levels, OEHHA can more clearly indicate that they can be used as an important reality check for selection of exposure-response models:

"The ethylene oxide cancer potency estimate derived from the NIOSH epidemiological study (see Section "Estimation of Cancer Potency" of this document) is based on excess risk. In other words, the human CSF expresses risk over and above the background risk. The background risk includes cancer risk due to endogenous exposures to ethylene oxide. Thus, in the case of ethylene oxide, the CSF is meant for use in computing risk levels associated with non-zero exogenous exposures (i.e., ambient air concentrations > 0 ppm).

This is a true statement for both the TCEQ and IRIS cancer risk calculation. This statement should not be used as a basis to ignore considering endogenous levels as part of an important reality check for derivation of cancer risk specific concentrations. It does not make sense for risk specific concentrations to be orders of magnitude below human endogenous levels, or to be a fraction of the population variability of human endogenous levels.

OEHHA's Proposed NSRL (p. 20) appears to incorrectly suggest that the dose-response relationship for endogenous ethylene oxide exposures is unknown at lower exposures:

"The dose-response relationship for endogenous ethylene oxide exposures within the homeostatic range might be different from the dose-responses seen with ambient exposures, possibly sublinear but ultimately unknown."

- We maintain that EtO hemoglobin adducts (HEV) are useful biomarkers of exposure, a point also noted in this publication, "*...Hb alkylation may serve as a particularly sensitive marker of exposure...*". It serves as an excellent cumulative measure of the internal doses of EtO present in blood for several months prior to measurement.
- Kirman analysis utilizes HEV to apportion total exposures to EtO from different pathways, which is a valid use of a biomarker of exposure.
- A toxicokinetic model is not required to utilize HEV data. Steady state blood levels of EtO (area under the curve or AUC) can be estimated from measured HEV levels with a high degree of confidence since the values for the reaction rate constants for EtO binding to hemoglobin and erythrocyte lifespans are known, as described in Motwani and Tornqvist (2014; see equations 2-3b).
- Although a toxicokinetic model that fully encodes the endogenous formation pathways (Kirman et al. 2021, Figure 4) are not yet available, the model of Filser and Klein (2018) and data cited therein (Figure 12A of Filser and Klein, 2018) are fully consistent with the use of a linear correlation between EtO in air and HEV measurements in humans, as adopted in Kirman et al. (2021) :

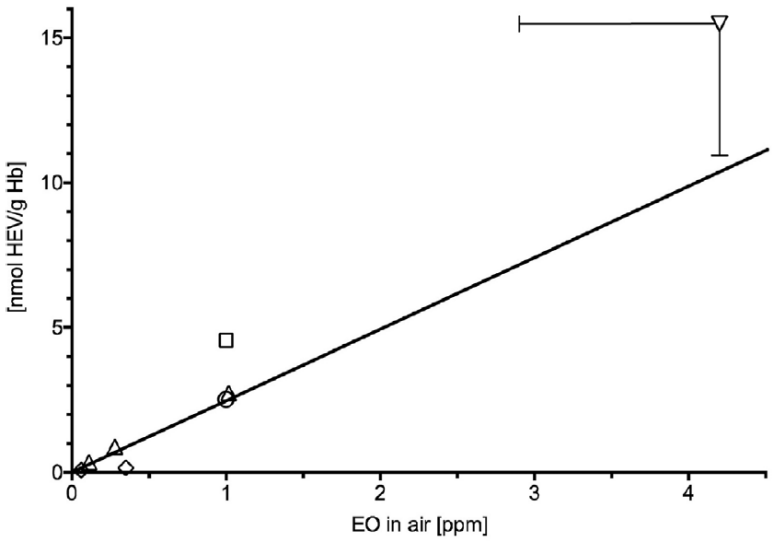

Figure 9. Hemoglobin adducts (HEV, nmol/g Hb) in workers exposed to EtO in air (from Filser and Klein, 2018).

- The available worker data depicted in this figure indicate that a linear relationship between HEV adducts and EtO in air is maintain across a broad range of concentrations (~0.1 ppm to ~4 ppm). Furthermore, the PBPK model of Filser and Klein (2018) predict a linear relationship across this range of exposures, as well as for exposures extrapolated below this concentration range (solid black line. Lastly, for exposure levels below the range of worker exposures, the NHANES biomonitoring data in smokers and in non-smokers are also consistent with a linear relationship (depicted by redline; note - log-linear scale) between EtO exposure (using cigarettes per day as a metric) and HEV adduct formation:

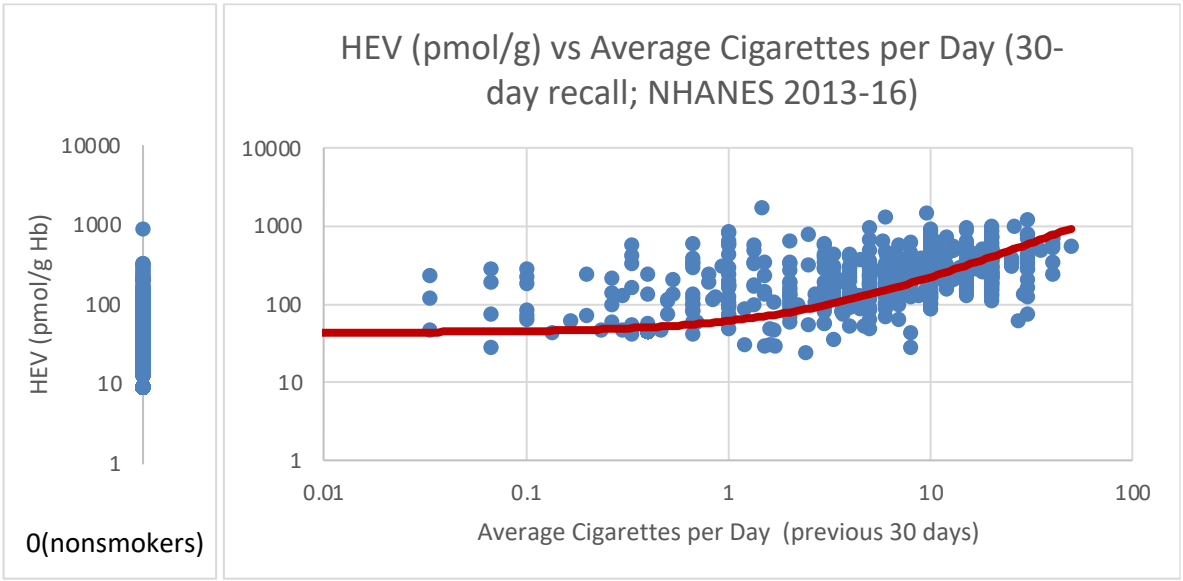

Figure 10. Hemoglobin adducts (HEV, pmol/g Hb) in U.S. smokers exposed to EtO in cigarette smoke (NHANES, 2013-16)

- These data indicate that the linear relationship between EtO exposure is maintained from background exposure levels up to 30x background levels. The linear relationship between cigarettes per day and HEV in this figure is consistent with a linear relationship (e.g., assessed by multilinear regression analyses) for another EO biomarker (urinary 2-hydroxyethyl mercapturic acid) as reported by CDC scientist (Kenwood et al. 2021). Together, these data provide strong and convincing evidence to support a linear relationship between HEV and EO exposure as used in Kirman and Hays (2017) and Kirman et al. (2021).
- Because the exogenous exposures to ET and EtO can be characterized with a high degree of confidence based upon available air monitoring data, and because there is high confidence in the NHANES biomonitoring data for HEV as a measure of total exposure to EO, estimates of endogenous exposure to EtO from these data can also be inferred with a high degree of confidence. There are no other known sources of EtO exposure that could contribute to the HEV levels measured by CDC.
- Measured HEV levels in human populations are dependent upon three parameters: (1) circulating levels of EtO in blood; (2) hemoglobin binding rates; and (3) erythrocyte lifespan. Although there are some possible sources of high-dose nonlinearity (e.g., induction of endogenous production of ethylene by EtO at high doses; theoretical exposure-related effects on erythrocyte turnover due to cytotoxicity) there are no documented sources of low-dose nonlinearity for EtO exposure and HEV formation. Any such nonlinearity would only be relevant at exposures above the range of observation (~0.1 to ~4 ppm) defined by Figure 12A above from Filser and Klein (2018). As noted above, there is no evidence of nonlinearity in the NHANES HEV biomonitoring data as a function of EO exposure (using cigarettes per day). As such this comment is inconsistent with available data sets, and inconsistent with default assumptions for chemical toxicokinetics.

The above comments show that there are reliable data for characterizing background endogenous and endogenous equivalent EtO exposure and provide a preface for using background exposure as context for managing and communicating EtO risk.

Kirman et al. (2021) model of external EtO exposures and internal EtO HEV hemoglobin adducts (EtOHEV) is validated in a "forward" analysis as suggested by EPA (2022)

The relationship between NHANES HEV biomonitoring data as a function of EO exposure (using cigarettes per day) established by Kirman and Hays (2017) and Kirman et al. (2021) is validated with a "forward" analysis, as suggested by EPA (2022), based measured EtO concentrations in mainstream cigarette smoke. Using the linear relationship between external EtO exposures and internal EtO HEV hemoglobin adducts (EtOHEV), Kirman et al. (2021) calculated that an approximate 10-fold increase in general population EtOHEV adducts in smokers compared to non-smokers (CDC NHANES, 2019) was equivalent to a continuous EtO air exposure of 21.7 ± 20.2 ppb (mean ± SD). EPA (2022) suggested that the Kirman exposure model could be validated if "forwards determinations of smokers total exposures to EtO" compared reasonably

to “backward” estimated EtO exposures derived from the Kirman EOHEV adduct/exposure relationship:

> p.69: “EPA also notes that the assumed relationship between HEV adduct measurements and EtO exposures in smokers (Kirman et al. 2017 and 2021) also needs validation. Cigarette smoke contains EtO and ethylene which may be metabolized to EtO. Smokers also experience physiological and biochemical changes that could affect their EtO exposures and/or formation of protein adducts. For validation of the HEV based projections, “forwards” determinations of smokers total exposures to EtO (e.g., as might be assessed using exhaled breath measurements) could be compared with “backwards” calculations of projected EtO exposure levels hypothesized from HEV from adduct level. Paired measurements of breath levels of EtO and ethylene and HEV adduct levels could provide useful bottom-line data to test the HEV/equivalent inhaled concentration hypothesis.”

Importantly, and directly responsive to the EPA-recommended validation exercise, multiple datasets have been published that describe reliable analytically-determined concentrations of EtO in individual cigarettes that can then be converted to total daily EtO smoker exposures dependent on the intensity of smoking behavior (Table 1; Liu et al, 2014; Forster et al., 2018; Jaccard et al., 2019).

Liu et al. (2014) reported for Kentucky Reference 3R4F cigarettes mean concentrations of 8.37 µg EtO/cig under the International Organization for Standardization (ISO) smoking regimen and 26.03 µg EtO/cig under the “Health Canada intensive” (HCI) smoking regimen. Forster et al. (2018) reported for the updated Kentucky Reference 1R6F cigarettes mean concentrations of 17.2 µg EtO/cig (HCI) and 19.3 µg EtO/cig (HCI) for Kentucky Reference 3R4F cigarettes. Jaccard et al. (2019) reported for Kentucky Reference 1R6F cigarettes mean concentrations of 5.92 µg EtO/cig (ISO) and 17.3 µg EtO/cig (HCI). Jaccard et al. (2019) also reported for Kentucky Reference 3R4F cigarettes which yielded mean concentrations of 6.78 µg EtO/cig (ISO) and 19.2 µg EtO/cig (HCI).

Daily EtO exposure concentrations (EC) can be estimated as C × CpD / IR, where C is the reported EtO concentration per cigarette (µg/cig), CpD is the number of cigarettes smoked per day (cig/day), and IR is the daily inhalation rate (m^3/day). CpD conservatively assumed to be 17 cig/day based on the average number of cigarette smoked by daily smokers in 2005 as reported by CDC (2018), and IR is assumed to be 16 m^3/day based mean inhalation rates for adults aged > 16 yr (EPA 2011). The estimated ECs, shown in Table 1 below, ranged from 3.5 to 15 ppb. These estimates are generally consistent with Kirman et al. (2021) estimates of 21.7ppb for smokers, 1.9 ppb for non-smokers, which results in 19.8 ppb from smoking contribution, and confirms that HEV adducts can provide reliable estimates of EO exposure.

Table 7. Estimated daily ethylene oxide exposure concentrations based on measured ethylene oxide concentrations in mainstream smoke

Source of EtO Mainstream Smoke Concentration Data	Reference Cigarette	Regimen	Estimated Daily Exposure Concentration (ppb)
	3R4F	ISO	4.98
Liu et al. 2014	3R4F	HCI	15.49
	1R6F	HCI	10.23
Forster et al. 2018	3R4F	HCI	11.48
	1R6F	ISO	3.52
	1R6F	HCI	10.29
	3R4F	ISO	4.03
Jaccard et al. 2019	3R4F	HCI	11.42

The data in Table 7 indicate that the "forward" analytical measurements of EtO in cigarette smoke, when converted to total daily EtO ppm exposures, are in excellent agreement with the "backwards" estimates of the mean and SD measurements of EtO ppm exposure calculated from the high-quality CDC smoker EtOHEV data using the Kirman EtO-EOHEV endogenous-equivalent model approach.

In addition, a preliminary analysis of the NHANES HEV data for smokers as a function CpD (see previous figure), demonstrates a linear relationship. The linear slope in this figure (18 pmol/g per average CpD) would correspond exactly with the slope of 10.9 pmol/g per ppb (continuous) if the conversion factor for ppb to CpD is approximately 0.6 ppb per CpD. Using the mean estimated daily concentration from Table 7 (8.9$\pm$4.3 ppb) along with the value of 17 CpD, results in an independently derived conversion factor of 0.53$\pm$0.25 ppb per CpD. Together these data indicate that the linear correlation between HEV in smoker exposures to EtO is excellent agreement with the linear correlation between HEV and occupational exposures to EtO (i.e., the slope of 10.9 pmol/g per ppb).

The consistency between the "forward" and "backward" smoking-derived EtO exposures can also be used to explore the plausibility of the IRIS IUR as a reasonable predictor of cancer risks associated with low EO exposures. If the IUR is assumed as correct, a 10 ppb (10,000 ppt, as a representative midpoint from Table 7) external EO exposure contributed by smoking is predicted to produce an upper-bound estimate on the order of $1x10^{-2}$ to 10^{-1} risk of cancers (i.e., the 0.1 ppt 10^{-6} risk projected by IUR) is 5 orders of magnitude less than approximate 10,000 ppt smoking exposures estimated by Kirman et al (2021) and validated by direct measurement of EtO in cigarettes). Such a conclusion suggests smoking should result in a readily demonstrable cancer signal in smokers where in fact the overall epidemiological data are weak or equivocal for this endpoint at best. Thus, the smoking data and associated EtO

exposure analyses are an important and reliable “reality check” that the IRIS IUR substantially overestimates the low-exposure cancer risks of EtO.

11. OEHHA cites two studies reporting an association between smoking and lymphohematopoietic (LH) cancers published in 2012 (Diver et al. 2012 and Kroll et al. 2012) to discount Kirman et al. (2021) reality checks. These two studies are inconsistent with the lymphoid cancer (NHL, lymphocytic leukemia, multiple myeloma) findings from the NIOSH mortality study (Steenland et al., 2004), upon which IRIS 2016 developed their low exposure high risk model and do not constitute a weight-of-evidence evaluation akin to the Surgeon General report, *The Health Consequences of Smoking —50 Years of Progress* (US DHHS 2014). In addition, OEHHA cites IARC review indicating a positive association between tobacco smoking and breast cancer, though not for lymphoid cancer.

** Pertaining to NSRL p. 19 **

OEHHA has responded to Kirman et al.'s plausibility argument that, if EtO caused lymphoid tumors, it would be seen in smoker studies, and such an association has not been reported. The Proposed NSRL (p. 19) cites two smoker studies published in 2012 (Diver et al. and Kroll et al.) in an attempt to provide biological plausibility for the IRIS cancer risk for lymphoid cancers:

> "*Since the IARC review, new results from two large prospective cohort studies have found significant associations with lymphoid cancer. The American Cancer Society Cancer Prevention Study II identified 1926 non-Hodgkin lymphoma cases in a cohort of 152,958 men and women (Diver et al., 2012). The study found an association between current smoking and non-Hodgkin lymphoma in women (RR = 1.37, 95% CI = 1.04–1.81), with a positive trend for years smoked ($p < 0.01$). The UK Million Women Study identified 7047 lymphoid cancers in a cohort of 1.3 million women (Kroll et al., 2012). This study found associations between tobacco smoking and Hodgkin lymphoma (1.45 per 10 cigarettes/day, 95% CI = 1.22–1.72) and mature T-cell malignancies (1.38 per 10 cigarettes/day, 95% CI = 1.10–1.73). These large-cohort findings support the plausibility of increased cancer risks from low concentrations of EtO.*"

These two studies do not support this statement. More importantly, they are inconsistent with the lymphoid cancer (NHL, lymphocytic leukemia, multiple myeloma) findings from the NIOSH mortality study (Steenland et al., 2004), upon which IRIS 2016 developed their low exposure high risk model.

Diver et al. is a large cohort study of the relationship between smoking among men and women and the risk for non-Hodgkin lymphoid neoplasms (NHL), a cancer endpoint in IRIS 2016, based on the NIOSH findings for this group of cancers. Diver et al. examined smoking history in detail including status, intensity, duration, cigarettes /day. OEHHA cites Diver et al.'s statistically significant RR for currently smoking women (1.37) but fails to note the deficit in currently smoking males (0.88), the positive trend with cigarettes per day and years smoked for females,

no trends for males. This lack of positive association in smoking males occurred despite the fact that males smoked more than females. Furthermore, they fail to note these gender increases are in the opposite direction of the NIOSH gender results associated with EtO exposure, SMRs=1.29 for males and 0.73 for females and the positive slope for males and the negative slope for females in internal analyses. Given the large number of cases in this study, they were able to demonstrate gender differences statistically. Despite gender differences, they also presented data for males and females combined. A statistically significant trend for NHL was seen for women but no trend for both genders combined. The results for other subtypes of the NIOSH lymphoid category either are positive for females only (lymphocytic leukemia) or non-positive for both genders (multiple myeloma). Diver et al conclude, "In the present study, current smoking was associated with an increased risk of NHL in women but not in men."

The Kroll et al. large cohort study was limited to females and examined the relationship between smoking and hematological cancers (both lymphoid and myeloid). OEHHA cites the increased risk observed in this study for **Hodgkin** lymphoma, which is not included in the NIOSH lymphoid category, which includes **non-Hodgkin** lymphoma. So OEHHA cites positive results but for a different disease. The authors do provide RRs for lymphoid cancers in their low exposure smokers (<15 cigarettes per day). No statistically significant increases were seen, and there was no evidence of steep increases at low concentrations.

Both papers summarize the existing literature related to smoking and hematologic cancers as "inconclusive" or "inconsistent", indicating a need for their research. Based upon their findings, Diver et al., conclude, "this large cohort study supports an association with cigarette smoking and increased risk of follicular lymphoma in women", while Kroll et al. conclude, "Cigarette smoking was associated with increased risk of Hodgkin lymphoma, consistent with previous reports." The weight of evidence related to smoking and lymphoid tumors (as defined by NIOSH) overall and at low exposure concentrations remain inconsistent and inconclusive.

EPA has also addressed this issue by raising the bar beyond a reality check of the existing literature requiring detailed quantitative analyses with adjustment for confounding and error bounds to rule out an association, ignoring theoretically easy detection of their putative IRIS conclusion of high risk at low exposures.

"As cigarette smoke contains many carcinogens, there is not a reason to expect, that EtO exposures to smokers would contribute a large part of total cancer risks due to cigarette smoking. A quantitative statistical analysis, which has not been reported, would be needed to place bounds on the potential levels of risk from lymphoid and breast cancers in smokers to support comparisons EtO cancer risks. Results from such analyses, appropriately controlled for other risk factors, might support reasonable comparisons of lymphoid and breast cancer rates in smokers and levels of risk of these tumors that would be predicted by EtO exposures from smoking. However, such analyses, to EPA's knowledge, do not appear to have been undertaken" (EPA 2022, p. 68).

The Agency concedes smokers have elevated EtO exposures “EPA notes that as smokers do have elevated exposures to EtO, further work to define and validate EtO exposure estimates and to determine statistical bounds on risks for EtO associated cancers in the smokers could in the future contribute important information for EtO risk assessment” (EPA 2022, p. 69).

Neither OEHHA nor EPA considered the published literature showing increased risk of acute myelogenous leukemia (AML) among smokers (IARC 2012b). The average smoker is exposed to 1.8 mg/day of benzene, which is ten times that of non-smokers (ATSDR 2007). These findings are plausible and a reasonable reality check, given that benzene is a known cause of AML. Extensive quantitative analyses as described above by EPA is not needed for this purpose, nor would it be needed to question the plausibility of an EtO/lymphoid tumor relationship based on highly exposed smokers.

The citation of these two studies reporting an association between smoking and LH cancers (Diver et al. 2012; Kroll et al. 2012) does not constitute a weight-of-evidence evaluation akin to the Surgeon General report, *The Health Consequences of Smoking —50 Years of Progress* (US DHHS 2014). More importantly, it does not address the point that an extraordinarily large potency estimate derived by USEPA for EtO is inconsistent with isolated or weak associations. HEV levels in smokers (236 pmol/g Hb; per NHANES) are equivalent to daily exposures to 780 ug/day, which is more than 4 orders of magnitude higher than the NSRL value, which would place predicted cancer risks above $1x10^{-1}$. Because smoking prevalence and intensity was much higher in the past, HEV burdens (and predicted risks from EtO exposures) would also be much higher than measured by NHANES. If the unit risk for EtO were truly predictive of its potency, the reported associations between smoking and lymphoid cancers would be larger and more consistently reported across epidemiological studies of smokers.

12. OEHHA's Proposed NSRL is an estimated EtO 10^{-5} risk-specific intake level. This intake level provides little utility in managing general population risk if background exogenous exposure isn't considered as an initial reality check. There has been an extensive ambient air measurement campaign over the last several years, including measurements near many sterilizer facilities and at background locations. In many cases, the levels of ethylene oxide far away from sterilizer facilities are similar to the levels near sterilizer facilities. Although the source makeup of this exogenous background ethylene oxide is currently not fully characterized, what is clear is that, in many cases, residents living near sterilizer facilities are not exposed to higher ethylene oxide than people living far away.

** Pertaining NSRL p. 49, 51 **

OEHHA's RSCs provide little utility in managing EtO risk

- OEHHA's proposed NSRL is based on the cancer unit risk of 3.3×10^{-3} per microgram per cubic meter $(\mu g/m^3)^{-1}$ from the EPA IRIS (2016a) assessment. Thus, based on the updated UR, an equivalent risk-specific concentration (RSC) for 10^{-5} cancer risk is 0.003 $\mu g/m^3$ or 0.002 ppb. This RSC is so low relative to background that it provides little utility in managing EtO risk.

EtO is quite unique among the managed hazardous substances

- Everyone in the U.S. is exposed to EtO regardless of where they live or work. Ethylene, the primary precursor of EtO, is released from natural and unregulated anthropogenic sources, abundant in ambient air, also produced metabolically and contributes substantially to background EtO exposure. EtO also is emitted from natural and likely unregulated anthropogenic sources and measurable in ambient air. The measurement of hemoglobin adduct 2-hydroxyethylvaline (HEV) has provided a biomarker of total background exposure to EtO, (exogenous EtO, exogenous ethylene and endogenous EtO). Background HEV levels have been measured in U.S. smokers and nonsmokers (CDC NHANES, 2019). The primary source of EtO background exposure in nonsmokers has been estimated to be endogenous production (reviewed by Kirman et al. 2021). These factors make EtO unique among most hazardous substances and indicate risk management challenges unless background exposure is considered.

- Ethylene is emitted to air from natural sources including plants where it functions as a hormone, microbial activity in soils, sediment and plant litter, as well as plants in aquatic systems (reviewed by Sawada and Totsuka, 1986; Morgott 2015; Health Canada 2016). A vast majority of ethylene emissions are from natural sources (Health Canada, 2016). Ethylene also is emitted from anthropogenic sources such as biomass burning, including

forest fires, and from exhaust emissions from gasoline and diesel vehicles (Swada and Totsuka, 1986; Margott, 2015; Health Canada, 2016). Early reviews also suggested EtO is emitted from vehicle exhaust (EPA 1985) although there is little current published supporting data. As both ethylene and EtO are constituents of tobacco smoke, combustion of tobacco (see Kirman et al., 2021), it seems reasonable to hypothesize that both ethylene and EtO may be emitted by biomass combustion and combustion exhaust from vehicles. As early as the 1970's, ethylene from natural and unregulated sources was being measured at relatively high levels in urban air (39-700 ppb; Ables and Heggestad, 1973). More recently, EtO has also been measured in ambient air at multiple locations across the U.S. away from known industrial sources to characterize this background exogenous source (median levels 0.03-0.33 ppb; ATSDR 2022; Sheehan et al. 2021; Lewis et al., 2022; Georgia EPD 2022).

- Because EtO is unique with ethylene as a precursor and its primary metabolic contribution to background exposure as well as EtO natural sources and likely unregulated anthropogenic source contributions to exposure, it is unlike nearly all other regulated hazardous substances and as such, presents unprecedented risk management and risk communication challenges. Considering that EtO management in California will be driven by the OEHHA risk assessment with its identified limitations, a reality check of OEHHA's proposed NSRL considering general population background exposure is warranted to better inform risk management of industrial emissions.

Based on the limitations of its risk assessment, OEHHA should consider including discussion of EtO general population background exposure to provide context for risk management and risk communication

OEHHA makes the following statements regarding background exposure,

> "*Measurements of specific hemoglobin adduct levels, such as N-2-hydroxyethylvaline (HEV), in humans or other species, reflect the integrated exposure to ethylene (endogenous + exogenous) and EtO (endogenous + exogenous). Kirman et al. (2021) showed background exposures to EtO and ethylene in ambient air alone are insufficient to account for HEV levels seen in non-smokers, and endogenous EtO production contributes more to non-smoker HEV levels than ambient EtO and ethylene exposures do. The EtO exposures from ambient and endogenous sources contribute to HEV levels, other adduct levels, and cumulative cancer risks (i.e., including from other chemicals and conditions). Thus, EtO and ethylene exposures are part of the risk factors accounting for the background cancer risk in the general population, including lymphoid and breast cancers (US EPA, 2016a; 2016b).*"

EtO background exposure contributes more to understanding risk than just accounting for background cancer risk; it provides a reality check on managing EtO general population risk when there are unaddressed questions about the representativeness of risk model. It is likely reasonable to assume that ambient exogenous and endogenous exposures have remained relatively constant over time, and therefore these exposures do not impact the conversions between relative risk and extra risk measures in the dose-response assessment. However, this assumption is not valid for exposures to EtO from smoking, which has changed significantly over time (Kirman et al., 2021) and can vary between worker sub-populations (e.g., salaried vs. hourly workers; Hsu et al., 2019). As such, smoking is a potential confounder for EtO exposures (vs. a confounder to observed cancer response) in the NIOSH cohort. In addition, there is a clear need to include discussion of background exposures for the purposes of risk management and risk communication of total EtO exposures and potential risks. Based on the limitations in the OEHHA EtO risk assessment described above and the unique characteristics of EtO background exposure, a risk management check based on general population background exposure is warranted as a reality check of the utility of OEHHA's assessment RSCs in managing general population EtO risk.

The draft OEHHA proposed NSRL provides little utility in managing general population risk if background exogenous exposure isn't considered as an initial reality check

The proposed OEHHA NSRL of 0.058 µg/day is equivalent to an inhalation 10^{-5} RSC of 0.003 µg/day or 0.002 ppb. For comparison purposes, the RSC inhalation exposure units are used. The RSC is a small fraction of substantially higher ambient background concentrations from natural and unregulated anthropogenic EtO sources otherwise associated with industrial emissions. The EtO technical support document reported ambient background EtO concentrations for the Los Angeles area ranging from 0.02-0.17 ppb. Based on data from the EPA national air toxics trends and urban air toxics monitoring programs for the October 2018 to September 2019 period, ATSDR estimated a national mean background EtO concentration of 0.13 ppb (ATSDR 2022).

Similarly, the same monitoring data for years 2018-2021 showed median (50^{th} percentile) background EtO concentrations ranged from 0.03-0.33 ppb nationally (summarized in Lewis et al. 2022). Local/regional location background EtO concentration for eight sterilization facilities evaluated over an extended period again showed median and 90^{th} percentile background concentrations ranging from 0.07 and 0.26 to 0.13 and 0.56 ppb, respectively (Sheehan et al. 2021). These data show that background concentrations are variable but more importantly, that everyone nationally (including Californians) is exposed exogenously to mean/median background concentrations of EtO substantially greater (~50-fold) than the proposed EtO RSC of 0.002 ppb.

Although it has been suggested that true background concentration based on a refined sampling method (TO 15A) may be lower than measurements based on EPA Method (TO 15), calculations from recent Georgia EPD background samples by both methods show refined background levels are still substantially greater than the OEHHA RSC.

An unwillingness of regulators to consider background exogenous exposure concentrations relative to RSCs as an initial reality check, particularly if concentrations at near facility locations are indistinguishable from concentrations at background locations, will face a serious risk management and communication problems (i.e., everyone in California is exposed to background EtO in ambient air from non-industrial background sources well above the RSC). As background EtO source emissions are not affected by managing industrial EtO emissions, managing risk under the proposed OEHHA risk assessment is untenable without considering ambient background levels.

The OEHHA updated EtO NSRL provides little utility in assessing general population risk unless total background exposure from combined exogenous and endogenous contributions are considered as an ultimate reality check

At some monitoring locations, there may be EtO concentrations significantly greater than background concentrations, or modeling may predict risk above the RSCs. The total exposure concentration compared to total equivalent background exposure concentration distributions for nonsmokers provides a final reality check of the utility of the OEHHA RSC and related NSRL estimates. There are published examples of where total equivalent concentration comparisons have been useful in informing whether further risk mitigation beyond recent emission controls was warranted.

For example, Sheehan et al. (2021) compared 50th and 90th percentile ambient exogenous concentrations from monitoring around eight facilities plus 50th percentile endogenous equivalent concentration (total exposure concentrations) with 50th and 95th percentile total equivalent background concentrations for nonsmokers in the U.S. population and concluded that facility concentrations are contributing negligibly to near residential population total exposure. Similarly, Lewis et al. (2022) compared 5-year average EtO modeled concentrations at near facility residences in Georgia plus background and endogenous concentrations (total exposure) with the 50th 60th and 95th total equivalent concentrations for the nonsmoking U.S. population and again concluded that facility contributions to residential exposure are negligible (see Figure 4 below from Lewis et al. 2022). These total exposure comparisons provide an additional reality check on the health significance of the facility emission contributions to near facility residential EtO exposure.

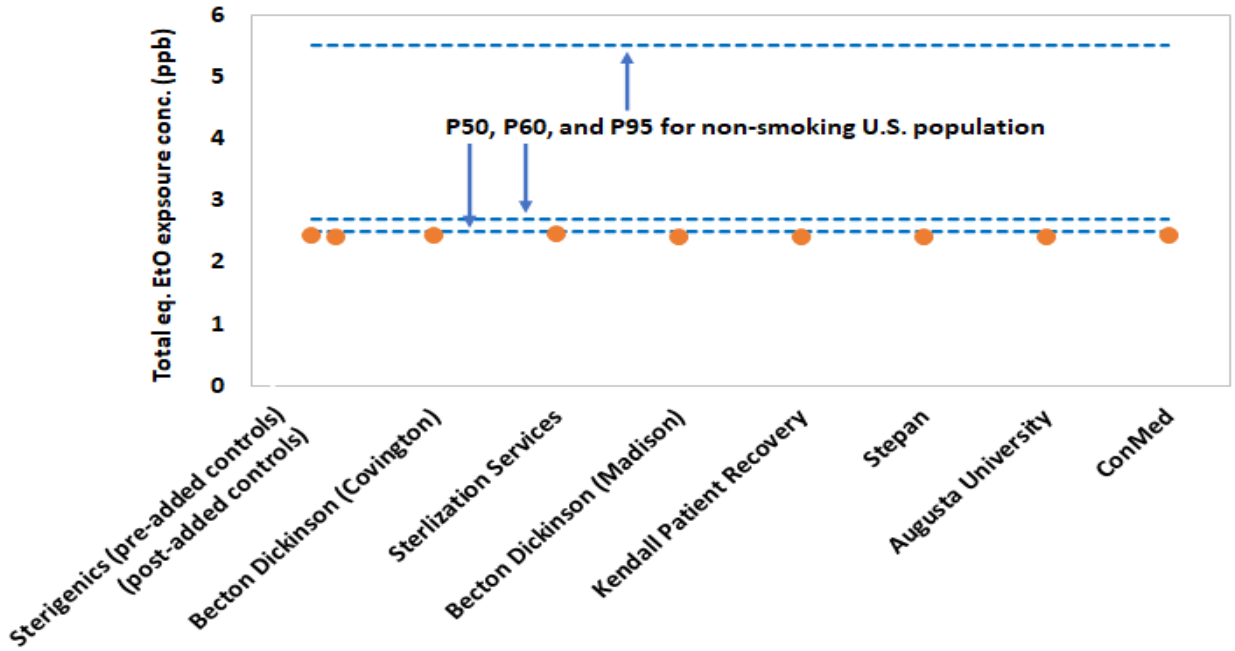

Figure 11. Estimated total equivalent exposure for the highest 5-year average modeled EtO concentration for all residential receptors by facility ((50th percentile endogenous equivalent for the non-smoking U.S. population, or 2.3 ppb) + (50th percentile background EtO concentration for Georgia, or 0.12 ppb) + (highest 5-year average modeled EtO concentration for all residential receptors by facility)) relative to that of the 50th, 60th, and 95th percentiles of the non-smoking U.S. population (2.5, 2.7, and 5.5 ppb, respectively).

We suggest that for populations in California living close to emitting facilities, OEHHA consider total equivalent exposure concentrations or the TCEQ RSCs as a final check in managing risk as the proposed OEHHA RSCs have no practical risk management utility.

We urge OEHHA to consider these comments and adopt an alternative such as the TCEQ risk value to derive the NSRL. Thank you.

Sincerely,

William Gulledge

William Gulledge

Senior Director

Chemical Products & Technology Division

References

Ables FB, Heggestad HE. 1973. Ethylene: An urban air pollutant. JAPCA 23:6 517-521.

American Conference of Governmental Industrial Hygienists (ACGIH). 1948. Documentation of the TLVs and BEIs: Ethylene Oxide; American Conference of Governmental Industrial Hygienists: Cincinnati, OH, USA, 2015.

American Conference of Governmental Industrial Hygienists (ACGIH). 1957. Documentation of the TLVs and BEIs: Ethylene Oxide; American Conference of Governmental Industrial Hygienists: Cincinnati, OH, USA, 2015.

Allison PD, 2010. Survival Analysis Using SAS®: A Practical Guide, Second Edition. SAS Institute Inc., Cary, NC.

Berman NG, Wong WK, Bhasin S, Ipp E. 1996. Applications of segmented regression models for biomedical studies. Am J Physiol 270(4 Pt 1):E723-32.

Bogen KR, Sheehan PM, Valdez-Flores C, Li AA. 2019. Reevaluation of historical exposures to ethylene oxide among U.S. sterilization workers in the National Institute of Occupational Safety and Health (NIOSH) Study Cohort. Int J Environ Res Public Health 16: 1738.

Boobis AR, Cohen SM, Dellarco V, McGregor D, Meek ME, Vickers C, Willcocks D, Farland W. 2006. IPCS framework for analyzing the relevance of a cancer mode of action for humans. Crit Rev Toxicol *36*(10): 781–792. https://doi.org/10.1080/10408440600977677

Bruch CW. 1961. Gaseous sterilization. Ann Rev Microbiol 15: 245–262.

Bruch CW. 1972. Sterilization of plastics: Toxicity of ethylene oxide residues. In Industrial Sterilization, Proceedings of the International Symposium, Amsterdam, The Netherlands, 2–4 May 1972; Phillips, G.B., Miller, W.S., Eds.; Duke University Press: Durham, NC, USA; pp. 49–78.

Bulka C, Nastoupil LJ, Koff JL, Bernal-Mizrachi L, Ward KC, Williams JN, Bayakly AR, Switchenko JM, Waller LA, Flowers CR. 2016. Relations between residential proximity to EPA-designated toxic release sites and diffuse large B-cell lymphoma incidence. South Med J 109(10):606-614.

Buonicore AJ, Desai PR, Mangone MA. 1984. Post Evacuation Cycle Modifications to Reduce Ethylene Oxide Residual Levels and Worker Exposure. Presented at the Medical Device and Diagnostic Industries (MD & DI) East, New York, NY, USA, 14–17; Chemrox Inc.: Bridgeport, CT, USA.

Burnham KP, Anderson DR. 2002. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach, 2nd ed. Springer-Verlag, New York, NY.

Carlsson H, Aasa J, Kotova N, Vare D, Sousa PFM, Rydberg P, Abramsson-Zetterberg L, Törnqvist M. 2017. Adductomic Screening of Hemoglobin Adducts and Monitoring of Micronuclei in School-Age Children. Chem Res Toxicol 30(5):1157-1167.

Diver WR, Patel AV, Thun MJ, Teras LR, Gapstur SM. 2012. The association between cigarette smoking and non-Hodgkin lymphoid neoplasms in a large US cohort study. Cancer Causes Control 23(8):1231–40. doi: 10.1007/s10552-012-0001-3. Epub 2012 Jun 12.

Filser JG, Klein D. 2018. A physiologically based toxicokinetic model for inhaled ethylene and ethylene oxide in mouse, rat, and human. Toxicol Lett 286: 54–79.

Fox AJ, Collier PF. 1976. Low mortality rates in industrial cohort studies due to selection for work and survival in the industry. Br J Prev Soc Med 30(4):225-30.

Garcia E, Hurley S, Nelson DO, Hertz A, Reynolds P. 2015. Hazardous air pollutants and breast cancer risk in California teachers: a cohort study. Environ Health 14:14.

Georgia Environment Protection Division (Georgia EPD). 2022. Ethylene Oxide Monitoring Report. Air Monitoring Branch, May 12, 2022.

Gkioulekas I, Papageorgiou LG. 2018. Piecewise regression through the Akaike Information Criterion using mathematical programming. IFAC Papers OnLine 51-15:730–35.

Goldgraben R, Zank N. 1981. Mitigation of Worker Exposure to Ethylene Oxide; Report Prepared for the U.S. Environmental Protection Agency; MTR-80 W333; Appendix C-10: Practices Relevant to Worker Exposure at Spice Industry Sites; The Mitre Corp.: McLean, VA, USA, March; pp. C153–C175.

Gollapudi BB, Su S, Li AA, Johnson GE, Reiss R, Albertini RJ. 2020. Genotoxicity as a toxicologically relevant endpoint to inform risk assessment: A case study with ethylene oxide. *Environ Mol Mutagen* 61(9): 852–871. https://doi.org/10.1002/em.22408

Greenberg HL, Ott MG, Shore RE. 1990. Men assigned to ethylene oxide production or other ethylene oxide related chemical manufacturing: a mortality study. Br J Ind Med 47(4):221-30.

Gridley G, Nyren O, Dosemeci M, Moradi T, Adami HO, Carroll L, Zahm SH. 1999. Is there a healthy worker effect for cancer incidence among women in Sweden? Am J Ind Med 36(1):193-9.

Hart JE, Bertrand KA, DuPre N, James P, Vieira VM, VoPham T, Mittleman MR, Tamimi RM, Laden F. 2018. Exposure to hazardous air pollutants and risk of incident breast cancer in the Nurses' Health Study II. Environ Health 17(1): 28. 10.1186/s12940-018-0372-3.

Health Canada. 2016. Screening Assessment Ethene (Ethylene) Chemical Abstracts Service Registry Number (CAS RN) 74-85-1; Health Canada: Ontario, CA, USA.

Hill AB. 1965. The environment and disease: Association or causation? Proc R Soc Med 58(5): 295–300.

Hornung RW, Greife AL, Stayner LT, Steenland NK, Herrick RF, Elliott LJ, Ringenburg VL, Morawetz J. 1994. Statistical model for prediction of retrospective exposure to ethylene oxide in an occupational mortality study. Am J Ind Med 25: 825–836.

Hsu CD, Momin F, Hess JW, de Carvalho MF. 2019. Trends in Cigarette Smoking Prevalence Among Refinery and Petrochemical Plant Workers, 1950 to 1999. J Occup Environ Med*61*(12): 989–995. https://doi.org/10.1097/JOM.0000000000001714

International Agency for Research on Cancer (IARC). 1987. Statistical Methods in Cancer Research. Volume II. The design and analysis of cohort studies. International Agency for Research on Cancer. Lyon, France.

International Agency for Research on Cancer (IARC). 1999. Cancer Epidemiology: Principles and Methods. World Health Organization. International Agency for Research on Cancer. Lyon, France.

International Agency for Research on Cancer (IARC). 2012a. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans. Chemical Agents and Related Occupations. Volume 100F. A Review of Human Carcinogens. World Health Organization. International Agency for Research on Cancer. Lyon, France.

International Agency for Research on Cancer (IARC). 2012b. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans. A Review of Human Carcinogens: Personal Habits and Indoor Combustion. Volume 100E. World Health Organization. International Agency for Research on Cancer. Lyon, France.

Jaccard G, Djoko DT, Korneliou A, Stabbert R, Belushkin M, Esposito M. 2019. Mainstream smoke constituents and *in vitro* toxicity comparative analysis of 3R4F and 1R6F reference cigarettes. Toxicol Rep 6:222-231. doi: 10.1016/j.toxrep.2019.02.009. PMID: 30886823; PMCID: PMC6402302.

Jones RR, Fisher JA, Medgyesi DN, Buller ID, Liao LM, Gierach G, Ward MH, Silverman DT. 2023. Ethylene oxide emissions and incident breast cancer and non-Hodgkin lymphoma in a U.S. Cohort. J Natl Cancer Inst 115(4):405-412.

Kenwood BM, McLoughlin C, Zhang L, Zhu W, Bhandari D, De Jesús VR, Blount BC. 2021. Characterization of the association between cigarette smoking intensity and urinary concentrations of 2-hydroxyethyl mercapturic acid among exclusive cigarette smokers in the National Health and Nutrition Examination Survey (NHANES) 2011-2016. Biomarkers 26(7):656-664. doi: 10.1080/1354750X.2021.1970809. Epub 2021 Sep 2. PMID: 34409911; PMCID: PMC8517914

Kirkeleit J, Riise T, Bjørge T, Christiani DC. 2013. The healthy worker effect in cancer incidence studies. Am J Epidemiol 177(11):1218-24.

Kirman CR, Hays SM. 2017. Derivation of endogenous equivalent values to support risk assessment and risk management decisions for an endogenous carcinogen: Ethylene oxide. Regul Toxicol Pharmacol 91: 165–72.

Kirman CR, Li AA, Sheehan PJ, Bus JS, Lewis RC, Hays SM. 2021. Ethylene oxide review: Characterization of total exposure via endogenous and exogenous pathways and their implications to risk assessment and risk management. J Toxicol Environ Health Part B Crit Rev 24: 1–29.

Kroll ME, Murphy F, Pirie K, Reeves GK, Green J, Beral V. 2012. Million Women Study Collaborators. Alcohol drinking, tobacco smoking and subtypes of haematological malignancy in the UK Million Women Study. Br J Cancer 107(5):879–87. doi: 10.1038/bjc.2012.333. Epub 2012 Aug 9.

Lewis RC, Sheehan PJ, DesAutels CG, Watson HN, Kirman CR. 2022. Monitored and modeled ambient air concentrations of ethylene oxide: Contextualizing health risk for potentially exposed populations in Georgia. Int J Environ Res Public Health 19: 3364.

Li W, He C, Freudenberg J. 2011. A mathematical framework for examining whether a minimum number of chiasmata is required per metacentric chromosome or chromosome arm in human. Genomics 97(3):186-92.

Manjanatha MG, Shelton SD, Chen Y, Parsons BL, Myers MB, McKim KL, Gollapudi BB, Moore NP, Haber LT, Allen B, Moore MM. 2017. Dose and Temporal Evaluation of Ethylene Oxide-Induced Mutagenicity in the Lungs of Male Big Blue Mice Following Inhalation Exposure to Carcinogenic Concentrations. Environ Mol Mutagen 58:122-134.

Marsden DA, Jones DJ, Britton RG, Ognibene T, Ubick E, Johnson GE, Farmer PB, Brown K. 2009. Dose-response relationships for N7-(2-hydroxyethyl)guanine induced by low-dose [14C]ethylene oxide: evidence for a novel mechanism of endogenous adduct formation. Cancer Res 69(7):3052-9.

Marsh GM, Keeton KA, Riordan AS, Best EA and Benson SM. 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. 10.1007/s00420-019-01438-z.

Mikoczy Z, Tinnerberg H, Björk J, Albin M. 2011. Cancer incidence and mortality in Swedish sterilant workers exposed to ethylene oxide: Updated cohort study findings 1972–2006. Int J Environ Res Public Health 8(6): 2009–2019. 10.3390/ijerph8062009.

Molinari N, Daure JP, Durand JF. 2001. Regression splines for threshold selection in survival data analysis. Statist Med 20:237–47.

Monson RR. 1986. Observations on the healthy worker effect. J Occup Med 28(6):425-33.

Morgott, DA. 2015. Anthropogenic and biogenic sources of Ethylene and the potential for human exposure: A literature review. Chem Biol Interact 241: 10–22.

Motwani HV, Törnqvist M. 2014. In vivo doses of butadiene epoxides as estimated from in vitro enzyme kinetics by using cob(I)alamin and measured hemoglobin adducts: an inter-species extrapolation approach. Toxicol Appl Pharmacol 281(3):276-84. doi: 10.1016/j.taap.2014.10.011. Epub 2014 Nov 5. PMID: 25448046

National Research Council (NRC). 2007. Models in Environmental Regulatory Decision Making. Washington, DC: The National Academies Press. https://doi.org/10.17226/11972.

National Toxicology Program (NTP). (1987). Toxicology and carcinogenesis studies of ethylene oxide (CAS no 75-21-8) in B6C3F1 mice (inhalation studies). National Toxicology Program Technical Report Series 326: 1-114.

National Toxicology Program (NTP). 2019. Risk of Bias Tool. National Toxicology Program (NTP). Updated August 05, 2022. Retrieved October 13, 2021, from https://ntp.niehs.nih.gov/whatwestudy/assessments/noncancer/riskbias/index.html.

Park RM. 2020. Associations between exposure to ethylene oxide, job termination, and cause-specific mortality risk. Am J Ind Med 63(7): 577–588. 10.1002/ajim.23115.

Recio L, Donner M, Abernethy D, Pluta L, Steen AM, Wong BA, James A, Preston RJ. 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.

Rietjens I, Michael A, Bolt HM, Siméon B, Andrea H, Nils H, Christine K, Angela M, Gloria P, Daniel R, Natalie T, Gerhard E. 2022. The role of endogenous versus exogenous sources in the exposome of putative genotoxins and consequences for risk assessment. Arch Toxicol 96(5): 1297–1352. 10.1007/s00204-022-03242-0.

Rinsky RA, Ott G, Ward E, Greenberg H, Halperin W, Leet T. 1988. Study of mortality among chemical workers in the Kanawha Valley of West Virginia. Am J Ind Med 13(4): 429–438. https://doi.org/10.1002/ajim.4700130403

Sawada S, Totsuka T. 1986. Natural and anthropogenic sources and fate of atmospheric ethylene. Atmos Environ 20(5): 821-832.

Sheehan PJ, Lewis RC, Kirman CR, Watson HN, Winegar ED, Bus JS. 2021. Ethylene oxide exposure in U.S. populations residing near sterilization and other industrial facilities: Context based on endogenous and total equivalent concentration exposures. Int J Environ Res Public Health 18: 607.

Shy CM, Kleinbaum DG, Morgenstern H. 1978. The effect of misclassification of exposure status in epidemiological studies of air pollution health effects. Bull N Y Acad Med 54(11):1155-65.

Sielken RL and Valdez-Flores C. 2009. Life-table calculations of excess risk for incidence versus mortality: Ethylene oxide case study. Regul Toxicol Pharmacol 55:82-89.

Steenland K, Whelan E, Deddens J, Stayner L, Ward E. 2003. Ethylene oxide and breast cancer incidence in a cohort study of 7576 women (United States). Cancer Causes Control 14(6): 531–539. 10.1023/a:1024891529592.

Steenland K, Stayner L, Deddens J. 2004. Mortality analyses in a cohort of 18235 ethylene oxide exposed workers: Follow up extended from 1987 to 1998. Occup Environ Med 61(1): 2–7.

Steenland K, Stayner L, Griefe A, Halperin W, Hayes, R, Hornung R, Nowlin S. 1991. Mortality among workers exposed to ethylene oxide. N Engl J Med 324 (20): 1402-1407

Stetson JB, Whitbourne JE, Eastman C. 1976. Ethylene oxide degassing of rubber and plastic materials. Anesthesiol 44: 174–180.

Swaen GM, Burns C, Teta JM, Bodner K, Keenan D, Bodnar CM. 2009. Mortality study update of ethylene oxide workers in chemical manufacturing: A 15 year update. J Occup Environ Med 51(6): 714–723. 10.1097/JOM.0b013e3181a2ca20.

Teta MJ, Benson LO, Vitale JN. 1993. Mortality study of EO workers in chemical manufacturing: a 10 year update. Br J Ind Med 50(8):704-9.

Texas Commission on Environmental Quality (TCEQ). 2015. TCEQ 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.

Texas Commission on Environmental Quality (TCEQ). 2020a. Ethylene Oxide Carcinogenic Dose-Response Assessment, CAS Registry Number: 75-21-8. Texas Commission on Environmental Quality (TCEQ). https://www.tceq.texas.gov/toxicology/ethylene-oxide

Texas Commission on Environmental Quality (TCEQ). 2020b. Response to External Peer Review Comments Received on the Ethylene Oxide Development Support Document, CAS Registry Number: 75-21-8. Texas Commission on Environmental Quality (TCEQ). https://www.tceq.texas.gov/downloads/toxicology/dsd/comments/eto-rtc-peer-review.pdf

U.S. Agency for Toxic Substances and Disease Registry (ATSDR). 2007. Toxicological Profile for Benzene. United States Department of Health and Human Services. Agency for Toxic Substances and Disease Registry (ATSDR). https://www.atsdr.cdc.gov/ToxProfiles/tp3.pdf.

U.S. Agency for Toxic Substances and Disease Registry (ATSDR). 2022. Toxicological Profile for Ethylene Oxide. United States Department of Health and Human Services. Agency for Toxic Substances and Disease Registry (ATSDR). https://www.atsdr.cdc.gov/toxprofiles/tp137.pdf.

U.S. Centers for Disease Control and Prevention (CDC). 2018. Smoking is down, but almost 38 million American adults still smoke. https://www.cdc.gov/media/releases/2018/p0118-smoking-rates-declining.htm

U.S. Centers for Disease Control and Prevention (CDC). 2019. National Report on Human Exposure to Environmental Chemicals: Ethylene Oxide Hemoglobin Adducts. U.S. Centers for Disease Control and Prevention. Available online: https://www.cdc.gov/exposurereport/index.html (accessed on 31 January 2019).

U.S. Department of Health and Human Services (US DHHS). 2014. The Health Consequences of Smoking: 50 Years of Progress. A Report of the Surgeon General. Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health.

U.S. Environmental Protection Agency (EPA). 1985. Health Assessment Document for Ethylene Oxide. United States Environmental Protection Agency, Office of Health and Environmental Assessment, EPA 600/8-84/009F.

U.S. Environmental Protection Agency (EPA). 2005. Guidelines for Carcinogen Risk Assessment. United States Environmental Protection Agency, Risk Assessment Forum, Washington, DC.

U.S. Environmental Protection Agency (EPA). 2011. Exposure Factors Handbook: 2011 Edition. United States Environmental Protection Agency, Office of Research and Development, National Center for Environmental Assessment. Washington, DC.

U.S. Environmental Protection Agency (EPA). 2016a. 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. United States Environmental Protection Agency, Office of Research and Development, National Center for Environmental Assessment. Washington, DC. Retrieved August 12, 2022, from https://cfpub.epa.gov/ncea/iris_drafts/recordisplay.cfm?deid=329730

U.S. Environmental Protection Agency (EPA). 2016b. Evaluation of the Inhalation Carcinogenicity of Ethylene Oxide: Appendices (CASRN 75-21-8). In Support of Summary Information on the Integrated Risk Information System (IRIS). United States Environmental Protection Agency (US EPA). https://cfpub.epa.gov/si/si_public_file_download.cfm?p_download_id=529971&Lab=NCEA.

U.S. Environmental Protection Agency (EPA). 2018. National Air Toxics Assessment. United States Environmental Protection Agency. Retrieved February 22, 2023, from https://www.epa.gov/national-air-toxics-assessment

U.S. Environmental Protection Agency (EPA). 2022. Summary of Public Comments and Responses for the Reconsideration of the 2020 National Emission Standards for Hazardous Air Pollutants: Miscellaneous Organic Chemical Manufacturing Residual Risk and Technology

Review. U. S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Sector Policies and Programs Division (E-143-01), Research Triangle Park, North Carolina 27711.

U.S. Environmental Protection Agency (EPA). 2023. Toxics Release Inventory (TRI) Program. United States Environmental Protection Agency. Retrieved February 22, 2023, from https://www.epa.gov/toxics-release-inventory-tri-program

U.S. Environmental Protection Agency Science Advisory Board (SAB). 2007. Review of Office of Research and Development (ORD) draft assessment entitled "Evaluation of the carcinogenicity of ethylene oxide." Washington, DC: Science Advisory Board, U.S. Environmental Protection Agency.

U.S. Environmental Protection Agency Science Advisory Board (SAB). 2015. Science Advisory Board Review of the EPAs evaluation of the inhalation carcinogenicity of ethylene oxide: Revised external review draft - August 2014 [EPA Report]. (EPA-SAB-15-012). Washington, DC: U.S. Environmental Protection Agency, Science Advisory Board.

U.S. Food and Drug Administration (FDA). 1978. Ethylene Oxide, Ethylene Chlorohydrin, and Ethylene Glycol: Proposed Maximum Residue Limits and Maximum Levels of Exposure; 43 FR 27474; Food and Drug Administration: Washington, DC, USA, 1978.

Valdez-Flores C, Sielken RL Jr. 2013. Misinterpretation of categorical rate ratios and inappropriate exposure-response model fitting can lead to biased estimates of risk: ethylene oxide case study. Regul Toxicol Pharmacol 67(2):206-14.

Valdez-Flores C, Sielken RL, Teta MJ. 2010. Quantitative cancer risk assessment based on NIOSH and UCC epidemiological data for workers exposed to ethylene oxide. Regul Toxicol Pharmacol 56: 312-320.

Vergnes JS, Pritts IM. Effects of ethylene on micronucleus formation in the bone marrow of rats and mice following four weeks of inhalation exposure. Mutat Res. 1994 Jul;324(3):87-91. doi: 10.1016/0165-7992(94)90051-5. PMID: 7517513.

Vincent MJ, Kozal JS, Thompson WJ, Maier A, Dotson GS, Best EA, Mundt KA. 2019. Ethylene oxide: Cancer evidence integration and dose-response implications. Dose Response 17(4): 1559325819888317. 10.1177/1559325819888317. https://journals.sagepub.com/doi/pdf/10.1177/1559325819888317

Von Grote J, Hürlimann C, Scheringer M, Hungerbühler K. 2003. Reduction of occupational exposure to perchloroethylene and trichloroethylene in metal degreasing over the last 30 years: Influences of technology innovation and legislation. J Expo Anal Environ Epidemiol 13: 325–340.

Von GroteJ, Hürlimann C, Scheringer M, Hunger K. 2006. Assessing occupational exposure to perchloroethylene in dry cleaning. J Occup Environ Hyg 3: 606–619.

Walker VE, Fennell TR, Upton PB, Skopek TR, Prevost V, Shuker DE, Swenberg JA. 1992. Molecular dosimetry of ethylene oxide: formation and persistence of 7-(2-hydroxyethyl)guanine in DNA following repeated exposures of rats and mice. Cancer Res. 52(16):4328-34. PMID: 1643630.

Walker VE, Wu KY, Upton PB, Ranasinghe A, Scheller N, Cho MH, Vergnes JS, Skopek TR, Swenberg JA. 2000. Biomarkers of exposure and effect as indicators of potential carcinogenic risk arising from in vivo metabolism of ethylene to ethylene oxide. Carcinogenesis 21(9):1661-9. doi: 10.1093/carcin/21.9.1661. PMID: 10964097.

White JD. 1977. Standard aeration for gas-sterilized plastics. J Hyg 79: 225–232.

Zeljezic D, Mladinic M, Kopjar N, Radulovic AH. 2016. Evaluation of genome damage in subjects occupationally exposed to possible carcinogens. Toxicol Ind Health 32(9): 1570–1580. 10.1177/0748233714568478.