Attachment 1

EO Panel Comments on HON Proposal- July 2023

INTRODUCTION

The proposed HON rule states:

"*Additionally, in 2016, the EPA updated the IRIS inhalation URE for EtO. In the first step of the CAA section 112(f)(2) determination of risk acceptability for this rulemaking, the use of the updated 2016 EtO risk value resulted in the EPA identifying unacceptable residual cancer risk driven by EtO emissions from HON processes. Consequently, the proposed amendments to the HON also address the EPA review of additional control technologies, beyond those analyzed in the technology review conducted for the HON, focusing on emissions sources emitting EtO that contribute to unacceptable risk."*

The Proposed HON Rule also importantly indicates that its actions are intended to "…*provide an ample margin of safety to protect public health following the new [EtO] IRIS inhalation UREs*" (HON Rule, p.2509000).

As such, the Proposed HON Rule reliance on the 2016 IRIS assessment is on unsound footing in that its URE has been extensively challenged by commentors identifying significant technical errors and deficiencies in the risk assessment. As the Proposed HON Rule depends on technically accurate and biologically plausible measures of "unacceptable residual cancer risk" the following comments identify and clearly describe the specific unaddressed flaws in the IRIS assessment that produced risk specific concentrations (RSCs) so small relative to background exposure as **to provide no utility in managing the residual risk** posed by facility EtO emissions. These comments also provide support for alternative risk assessments, such as the TCEQ assessment that provides biologically plausible RSCs and health-based action levels to identify and manage unacceptable residual cancer risk under the HON amendments for EtO.

The lack of utility is particularly apparent in the HON proposal that sets an EtO Action Level based not on acceptable EtO health risk as estimated from the IRIS URE, but rather on using analytical detection limits appropriate for routine monitoring of EtO fenceline concentrations above and beyond its demonstrated natural presence in ambient air. The analytical detection limit based Proposed HON Action Level of 0.2 µg/m3 (0.1 ppb), that is defined as being over and above EtO naturally present in background ambient air, is equivalent to an approximately 10^{-3} cancer risk as estimated by the IRIS IUR (i.e., 0.0001 ppb = 10^{-6} cancer risk). This estimated 10^{-3} cancer risk projected for the Proposed HON Action Level exceeds the 1 in 10 thousand cancer risk EPA generally assumes as a "presumptive limit of acceptability" (HON, p.21111), and for which no further facility corrective actions are otherwise required. Thus, the Proposed HON Action Level that relies on the flawed IRIS IUR creates a distinct conundrum regarding communicating facility risks to affected communities, i.e., an inference to communities that

facilities in compliance with the proposed Action Level still present EPA-defined unacceptable cancer risks.

The Proposed HON itself indicates a reasonable science-based resolution to the lack of the utility of the proposed EtO Action Level. The Proposed HON states:

"For residual risk assessments, we [EPA] generally use UREs from the EPA's IRIS. For carcinogenic pollutants without IRIS values, we look to other reputable sources of cancer dose-response value, often using California EPA [CalEPA] UREs, where available. In cases ***where new scientifically credible dose-response values have been developed in a manner consistent with EPA guidelines and have undergone a peer review process similar to that used by the EPA, we may use such dose-response values in place of, or in addition too, other values*** [emphasis added],"

As has been extensively described in previous ACC comments (2020) and further detailed in this EO Panel response to the Proposed HON, the Texas Commission of Environment Quality (TCEQ, 2020) published a "new" (post 2016-IRIS) "scientifically credible" and peer-reviewed exposure-response model importantly based on the same occupational cohort as that relied on by IRIS for its URE. The TCEQ exposure-response model estimates an EtO URE that is approximately 3 orders of magnitude greater than that estimated by IRIS. Although EPA has challenged the TCEQ URE as a reasonable alternative cancer exposure-response model (insert EPA response to MON reference; EPA, 2022), previous ACC comments and the current ACC HON comments extensively detail how the TCEQ URE represents a higher-quality and more "scientifically credible" EtO cancer risk specific concentrations (RSCs) compared to the EPA IRIS RSCs.

Based on the TCEQ approach to use the standard CPH model, these comments provide useful and biologically plausible RSCs and basis for Action Levels for identifying unacceptable residual cancer risk and support the HON amendments for EtO (Detailed Comments Part 1). In addition, we provide risk management comments regarding how background exposure should be considered in deriving Action levels (Detailed Comments Part 2). We also note that the Clean Air Act does not require that EPA use any particular URE in making decisions under Section 112. Accordingly, if the Agency decided not to adopt the TCEQ value it could incorporate the information contained herein to recognize the problems with its IRIS value as significantly overstating risk and develop a path based on that understanding.

DETAILED COMMENTS: PART 1

EPA's (2022) response to comments include new information and analyses that are directly relevant to the proposed HON Rule. EPA (2022) often misrepresents selected criticisms of TCEQ peer reviewers; expands EPA's highly subjective "eyeballing" of exposure-response models to the biological dose-response data; and criticizes TCEQ or ACC's analyses incorrectly for issues that are far more problematic in the EPA (2016) analysis. Our comments below primarily address EPA's (2022) response to comments. A more complete response to EPA (2022) can be found in our earlier comments on the MON and in recent comments to OEHHA (ACC, 2023) that address new analyses and publications.

EPA (2022) separates comments into sub-topics. In doing so, some of ACC's comments are blurred in with other distinctly different public comments and there is a loss in the true meaning and emphasis of ACC's comments . Thus, our comments summarize the following key points made in earlier ACC comments followed by more detailed response to specific points made by EPA (2022).

1. The Steenland and IRIS analyses are based on extensive statistical modeling analyses absent any consideration of the epidemiological weight of evidence and biological plausibility.
    a. The emphasis on log cumulative models as the driving force for justifying a steep exposure response model based on visual fit and statistical analysis is inappropriate.
    b. The emphasis on internal categorical estimates based on combined males and females is inappropriate based on the Steenland et al. (2004) conclusions.
    c. Following EPA's lead for how it presented visual fit comparisons, TCEQ like EPA emphasized the pattern of the data that is modeled. EPA unfairly uses the absence of confidence intervals (also missing in EPA's graphs) as a means to disprove use of the CPH model. In fact, TCEQ is illustrating why visual fit using a few categorical rate ratios (RR) based on grouped data should not be a basis to determine goodness of fit of models based on individual data.

2. The epidemiological evidence does not support a steep exposure response. EPA dismisses evidence from the UCC study because it incorrectly concluded the NIOSH study has superior exposure data and criticizes the UCC study for long follow up ignoring the fact that there were no findings at any stage of analyses.

3. Breast cancer, like other types of cancers 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 quantitative risk assessment based on substantial under ascertainment of incident cases reported by Steenland et al (2003) and subsequent risk deficits in the lower exposures.

4. The biological evidence does not support a steep exposure response, and together with the epidemiological evidence described above, should be the major driver for selection of exposure response models

5. Both the CPH model and the IRIS model estimate extra risk, so this fact should not be used as a basis to ignore reality checks based on valid estimates of endogenous levels. 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 at and substantially below that the body produces through natural processes.

6. TCEQ's prediction analyses provides an important reality check and is based on well accepted methods also used by IRIS (2016). This approach is superior to the subjective visual fit methods. Calculation of confidence intervals by ACC using a different method from TCEQ's results in nearly identical CI's and further support conclusions based on TCEQ's prediction analyses.

1. **The Steenland and IRIS analyses is based on an extensive statistical modeling exercise absent any consideration of the epidemiological weight of evidence and biological plausibility. The emphasis on log cumulative models and categorical estimates for both sexes combined as the driving force for justifying a steep exposure response model is inappropriate.**

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. On the one hand, EPA states that their evidence integration to inform MOA for hazard assessment was adequate and appropriate. On the other hand, the agency argues that "not enough is known about the MOA to use evidence integration to inform the shape of a dose-response model representing cumulative lifetime human exposure." These contradictory positions precluded the integration of valuable information on biological plausibility as a major factor in selecting among different statistical models. Instead, the IRIS exposure-response assessment is driven by repetitive statistical modeling (Steenland et al., 2003, 2004; EPA, 2022). The EPA (2005) carcinogen risk assessment guidelines captures the general issue of applying multiple curve-fitting models based on statistical modeling approaches applied to the EPA IRIS (2016a) assessment for EtO:

> "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 lymphoid findings in the UCC study and in NIOSH females at any exposure, and absence of statistically significant findings at lower exposures in males in the NIOSH study are more consistent with a standard CPH model than an extremely steep initial exposure-response slope.

In response to ACC (2020) comments, EPA (2022) places great emphasis on the statistical significance of the log cumulative model as an important reason for dismissing ACC's and TCEQ's support for the TCEQ CPH model. At the same time, EPA downplays the importance of p-values and AIC when addressing comments on the incorrect calculation of p-values. ACC agrees that statistics cannot be the basis for selecting between EPA's and TCEQ's models because both models are not statistically significant for lymphoid cancers, and both models are statistically significant for breast cancer incidence when the statistical analyses are corrected.

Nevertheless, EPA's (2022) repeated emphasis on the statistical significance of the log cumulative model indicate that EPA (2022) places much more reliance on a model it has also rejected because this model is biologically implausible.

In considering the statistical significance of the log cumulative model, it is important to keep perspective that a large number of curve-fitting models were applied. The EPA IRIS statistical modeling builds on a backdrop of a large number (>50) of statistical models already conducted in the original published papers by Steenland et al. (2003, 2004). The NIOSH mortality study initially conducted CPH modeling using a variety of exposure metrics (duration, average, maximum, cumulative, and log cumulative exposure) and both continuous and categorical exposure variables. For each of these combinations, a variety of lags were also examined (no lag, 5, 10, 15 and 20) but the authors indicate that "in the results, we present only the lagged model with the best fit to the data, as judged by the likelihood ratio test". This analysis is not based on any *a priori"* hypothesis regarding biological mode of action, epidemiological evidence or toxicokinetics, and the IRIS assessment further expands the analyses to include a large number of additional models.

The only continuous dose-response models that were reported by Steenland et al. (2004) to have statistically significant positive slopes for the given metric were

(1) Lymphohematopoietic (LH) cancer mortality: males, log cumulative exposure model with 15 year lag
(2) Lymphoid cancer mortality: males, log cumulative exposure model, and 15 year lag
(3) Breast cancer mortality : log cumulative exposure model and 20 year lag.

The log cumulative exposure model was not significant for LH and lymphoid cancers for females alone or it would have been reported by Steenland et al. (2005). Thus, the evidence of an increased risk of LH, lymphoid and breast cancer mortality (p-values = 0.02, 0.02, and 0.01, respectively) were for the logarithm of lagged cumulative exposures to EO. The model estimated the slope and also estimated the lag using the likelihood of the model as the criterion to select the lag. Therefore, the lag and the slope are two parameters (two degrees of freedom as opposed to one) being estimated from the observed data and should be included in the evaluation of the statistical significance of the model. If Steenland (2004) had accounted for the lag as another degree of freedom, then the p-values would have been 0.07, 0.07, and 0.06, respectively. These resulting p-values indicate that the models are not statistically significantly different at the 5% significance level than a model with no exposure response (the null model).

EPA IRIS (2016) focused on log cumulative exposure models as a major driving force but did not take into account that these models force supralinearity that could be simply linear (Valdez-Flores et al., 2010). A case in point is that the breast cancer incidence data is statistically significant based on the log-linear CPH model, the linear CPH model, and the log cumulative CPH model. As noted above, EPA (2022) presentation of statistically significant models for breast cancer incidence omitted the fact that the log-linear CPH model for breast cancer

incidence is statistically significant, which makes it more difficult for the public to accurately compare the two models.

These log cumulative statistical modeling results from the NIOSH mortality study (Steenland et al. 2004), together with the shape of the internal categorical estimates combining males and females together led EPA to select the two-piece spline model with the knot at 1600 ppm-years based on statistical fit and visual comparisons. For the categorical analyses, Steenland et al (2004) reported statistically significant increases for the same three endpoints LH (males only), Lymphoid (males only) and Breast cancer mortality (females only) but only in the highest exposure category. Although the methods section of Steenland et al. (2004) clearly indicate the analyses were conducted with males only, females only, and both sexes, Steenland appropriately reported the results of males and females separately and concluded that the exposure response pattern in male workers was distinctly different from that of female workers. Thus, ACC emphasizes that the original authors' conclusions indicate that the epidemiological evidence supports the CPH model and not a 2-slope linear spline model with an initial steep slope followed by a shallow slope. ACC (2020) comments emphasized this issue, but EPA did not respond to the obvious inconsistencies between their preferred model and the conclusions of Steenland et al. (2003, 2004).

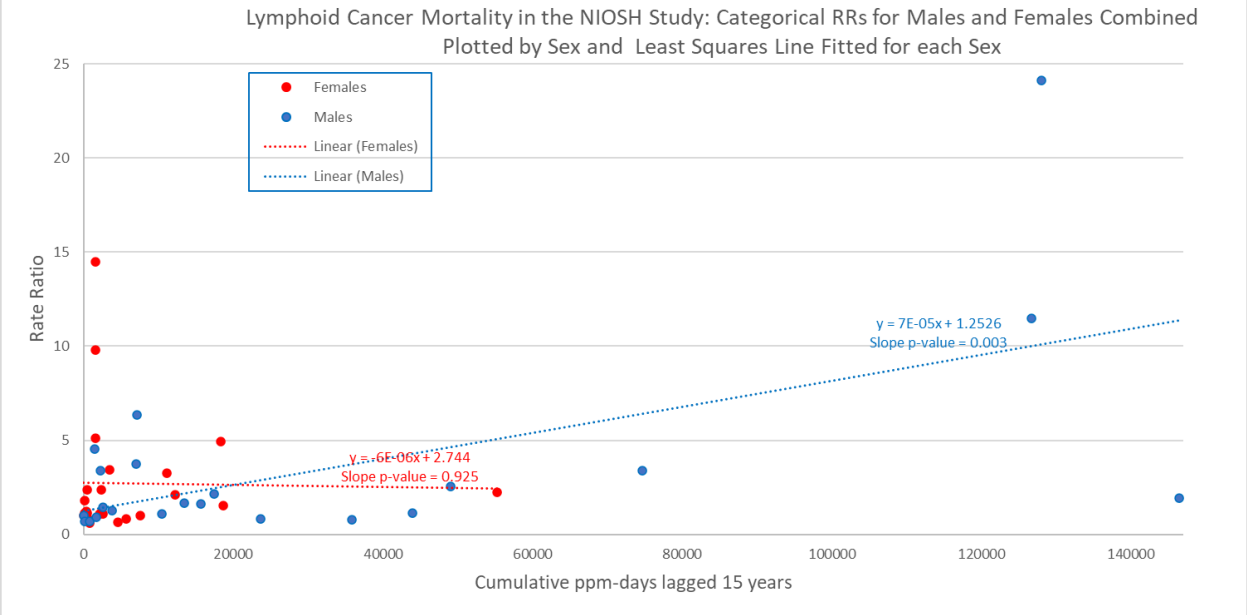
Briefly, Steenland et al. 2004 concluded, "Positive exposure response trends for lymphoid tumours were found for males only. Reasons for the sex specificity of this effect are unknown." In the Cox regression continuous cumulative exposure for lymphoid tumours (mortality), the authors report a positive slope for males and a negative slope for females. Similarly, in categorical cumulative exposure analyses, the same slope patterns prevailed. The non-statistically significant odds ratios for males were 1.00 2.45 1.85 2.44, while for females they were 1.00 2.05 1.25 0.87. The male- female differences were also apparent in the initial external analyses where the SMRs for NHL (the largest component of lymphoid mortality group) was 2.37 and statistically significant for males in the highest cumulative exposure category, compared to 0.37 based on one case in females. Given these obvious differences in both external and internal analyses, Steenland et al.'s regression results with log cumulative exposure and 15 yr. lag were presented in the paper for males only.

Despite these obvious differences, EPA in IRIS combined males and females and derived RRs based on exposure groupings of the pooled data. EPA's visual fit exercise used these categorical estimates of RRs as the "gold standard" and proceeded to graphically compare the shape of various models to the pattern of the 4 combined male/female RRs. The pattern suggested supralinearity to the Agency. However, the exposure-response pattern is not steep at low exposures when examining the male data alone. Specifically, the male RR for lymphoid mortality in the lowest exposure category (Table 7 of Steenland et al. 2004) was 0.90 (95%CI: 0.16-5.24). Since none of the RRs were statistically significantly different from 1.0, absent any other information to the contrary, one might argue in favor of combining the data. This approach, however, of relying on a p-value as opposed to the clear pattern of differences between males and females across several lines of analysis, flies in the face of obvious and more appropriate epidemiologic interpretation.

ACC (2020) provided figures that illustrate the very different pattern in the estimates of the categorical rate ratios supporting Steenland et al. (2004) conclusion that there was an increase in lymphoid mortality in the highest cumulative exposure category only in males but none in females. Figure 1 below shows the categorical estimates of the rate ratios for male workers (blue dots) and female workers (red dots) shown similarly to TCEQ (2020) that show this same pattern for each worker. A least square line was fit to the categorical rate ratios of male (blue line) and female (red line) workers separately. The line for the rate ratios of male workers (has a statistically significant positive slope (7×10^{-5} with p-value = 0.003), while the line for the rate ratios of female workers has a negative slope (-6×10^{-6} with p-value = 0.925) that is not statistically significantly different than zero.

The distinctly different pattern between females and males for lymphoid cancers demonstrates more clearly why steep exposure-response models based on both males and females combined do not accurately reflect the epidemiological evidence and conclusions made by Steenland et al. 2004. In addition, fitting log cumulative models to the pooled male and female data, results in a model that tries to accommodate the decreasing trend in females (with smaller cumulative exposures) and the increasing trend in males (with larger cumulative exposures), when there is clear epidemiological evidence that the data from both sexes should not be combined. Although only males should be modeled, any risk estimate for lymphoid cancer based on males will be protective of females, for whom there was no EtO/lymphoid cancer association .

FIGURE 1. PATTERN OF RATE RATIOS FOR FEMALE AND MALE WORKERS WITH TREND LINES USING LEAST SQUARE BUILT IN EXCEL



Source: Table D-34 in the Appendices to the IRIS 2016, EPA lists the cumulative exposure to EtO lagged 15 years for each of the lymphoid decedents in the NIOSH study. Table 9S in the supplementary material of Valdez-Flores and Sielken (2013) paper lists the categorical rate ratios estimated for the 9 (6 males and 3 females) lymphoid deaths in the unexposed workers and for each of the 44 lymphoid decedents with cumulative exposures to EtO lagged 15 years greater than 0. The categorical rate ratios reported in

Valdez-Flores and Sielken (2013) were plotted against the cumulative exposures to EtO in IRIS 2016. In addition, the lymphoid decedents were classified according to sex (male or female worker).

There are several inconsistencies and flaws with EPA's approach and reliance on statistical significance of the log cumulative exposure model and internal categorical estimates of RRs to drive the selection of a 2-slope linear spline model for quantitative exposure-response analysis based on statistical and visual fit criteria rather than consideration of the epidemiological evidence and biological plausibility. Although our focus is on the visual fit and statistical analyses of the lymphoid mortality data, some of these issues apply to the breast cancer incidence data as well. However, the CPH, spline and log-cumulative models were statistically significant for breast cancer incidence, and the breast cancer incidence data should not be used due to concerns about underascertainment.

ACC (2020) provided detailed comments regarding EPA IRIS (2016a) flawed statistical and visual fit justification which cannot be divorced from the epidemiological and biological evidence.

- First, EPA (IRIS, p. H-26, H-34) rejects the log cumulative exposure models because these models are "excessively sensitive to changes in exposure level in the low-dose region, and thus, were not biologically realistic". It is inappropriate to use the statistical significance of a biologically implausible exposure response model as a rationale to dismiss the TCEQ CPH model.

- Second, EPA's *a priori* criteria for selecting the two-piece spline model was to systematically evaluate different knots in increments of 100 ppm-days and selecting the knot that resulted in the best (largest) model likelihood. The statistical basis for selection of the spline model is flawed because the knot was clearly a parameter in the model. As discussed in detail in earlier ACC (2020) comments, the corrected p-values indicate that there is no statistical basis to select the spline model over the CPH model. For lymphoid mortality, the p-values are 0.14 (corrected from 0.07) and 0.22 for the spline and CPH models, respectively. For breast cancer incidence, the p values are 0.04 (corrected from 0.01) and 0.02 for the spline **and <u>also the CPH model</u>**. For both lymphoid and breast cancer incidence, the CPH model is a preferred more parsimonious model that has greater biological plausibility.

- Third, EPA (2022) introduces a new visual evaluation of the lymphoid categorical data to provide evidence that the CPH model is inappropriate because it explodes upwards at higher cumulative exposure. These analyses do not reflect the TCEQ exposure-response analysis

- Fourth, and as discussed in detail above, the categorical rate ratios for males and females combined were used for visual comparisons for the different models that IRIS applied

without any consideration of the original NIOSH authors' peer reviewed published conclusions that males have a distinctly different pattern than females.

- Fifth, the EPA (2022) criticisms of the TCEQ graphs of individual data miss the larger point that their purpose is to illustrate why the EPA IRIS figure notes are correct that comparisons should not be made along the y-axis. The EPA IRIS (2016) figures misled EPA and EPA SAB to make comparisons along the y-axis based on public comments received regarding underestimation of the "data." Of greater concern is that EPA (2022) incorrectly states that the continuous modeling of the data involves comparison against the same reference group, which perpetuate this misunderstanding that all models have the same implicitly estimated baseline risks, which would permit comparisons along the y-axis.

- Sixth, the EPA IRIS selection of models did not consider the putative genotoxic mode of action, pharmacokinetic data and other animal data in any integrated manner to inform the selection of the dose-response model. As described in greater detail below, the standard CPH model has greater biological plausibility than the spline model (addressed in detail in a later section).

EPA (2020) rationale that the steep log cumulative exposure model is statistically significant.

It is contradictory for EPA to use a model that EPA rejected (steep log cumulative exposure) to justify the selection of a spline model. We strongly agree with EPA that there are concerns with the log cumulative exposure model itself that are not necessarily EtO specific. All the models tested by Steenland (2011) are exercises in curve fitting. The log cumulative models are among the multiple models that can be used to analyze data that approach an asymptote, i.e., flatten out at higher exposure or dose levels. That one model fits better than another by a statistical criterion such as AIC does not make it more biologically realistic. Mathematically, the log cumulative forces a very steep initial slope that EPA (2016) recognized is biologically implausible. The log cumulative model, which is forced through RR=1 at zero exposure, describes the data well precisely because the exposure-response curve is essentially flat. The irony obviously is that the flatness of the exposure-response curve would lead many epidemiologists to conclude that there is no causal association here but leads the risk assessor to estimate an implausibly high risk at low exposure levels. Valdez-Flores et al. (2010) provides additional explanation of the statistical and mathematical issues with the log cumulative model that are independent of EtO.

EPA (2020) rationale that the knot is optional to consider as a parameter in the model.

EPA's (2022) response to ACC's comments on the error in the statistical significance is to dismiss the strong concurrence from TCEQ peer reviewers with statistical modeling expertise by stating EPA ". . .does not disagree that some modelers would follow an alternate process where the knot location is handled as a fully adjustable model parameter and uncertainty in the

parameter. . .If followed, this process would lead to some increase in the calculated fit statistic (AIC) and model p-values."

EPA (2022) does not present exactly what these two peer reviewers state:

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[sic14], 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)

What EPA (2022) dismissively describes as "some increase" is in fact a substantive difference between what EPA describes as nearly significant p=0.07 to a clearly non-significant p=0.14. Had both the statistics and stronger cautions in IRIS against visual fit comparisons been accurately presented in the EPA IRIS, EPA SAB likely would have reached very different conclusions. This is apparent in the EPA (2019) sensitivity analyses which selected models based on the incorrect p-values and visual fit criterion, as well as in public comments, including those by the former chemical manager of the IRIS (2016), that base their comments on these incorrect statistics and flawed comparisons of visual fit along the y-axis.

EPA (2022) then quotes EPA SAB in justifying the statistical treatment in relation to the knot: "The knot is preselected and is not considered a parameter in these analyses, consistent with SAB's concept of parsimony [footnote 14: "in some setting the principle of parsimony may suggest that the most informative analysis will rely upon fixing some parameters rather than estimating them from the data. The impact of the fixed parameter choices can be evaluated in sensitivity analyses. In the draft assessment, fixing the knot when estimating linear spline model fits from relative risk regressions is one such example [page 12 of SAB (2015)]."

However, prior to fitting its two-piece spline model, EPA did not simply "fix" or "select" the position of the knot in that model "rather than estimating" the knot position, as specified by the SAB. Instead, EPA tested multiple knots, and selected knot values that maximized the

likelihood of data fit. Just because the analyses were done in two steps, one performed with one program and another with a second program, does not mean the estimated parameter in the first step should be ignored in the p-value calculation.

If EPA (2022) reinterprets the SAB to indicate that the SAB agrees that the method EPA used to calculate the p-value does not need to include the statistical optimization of the knot based on maximizing the likelihood, then the SAB is incorrect. Just because the analyses were done in two steps, one performed with one program and another with a second program, does not mean all estimated parameters should be ignored. Only TCEQ (and not EPA) specifically asked peer reviewers the question of whether the knot should be included as a statistical parameter. As quoted above, the peer reviewers who clearly were the two statisticians on the expert panel agreed with TCEQ's correction of the EPA IRIS p-values.

ACC's (2020) comments also highlighted a basic principle to account for all model parameters for regulatory decision making as clearly stated by the National Research Council report "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). Thus, EPA's (2022) dismissal of the TCEQ peer reviewers agreement with TCEQ (2020) that p-values were incorrectly calculated, is inappropriate especially since these p-values played a major role in the IRIS (2016) assessment and EPA SAB's conclusions. This substantive correction should not be so lightly hand-waved away as an optional approach that is a matter of opinion or preference among statisticians.

Thus far, all that EPA has provided to support its approach is that it believes its approach to be equally valid, but if that is the case EPA should have been able to provide literature supporting its position, rather than dismissing the several publications on spline-modeling provided by ACC (2020). The overwhelming evidence presented by ACC (2020) and TCEQ (2020) is that the p-values for the spline model are incorrect and leads to a substantive change in the summary tables that EPA SAB (2015) reviewed (Table 1).

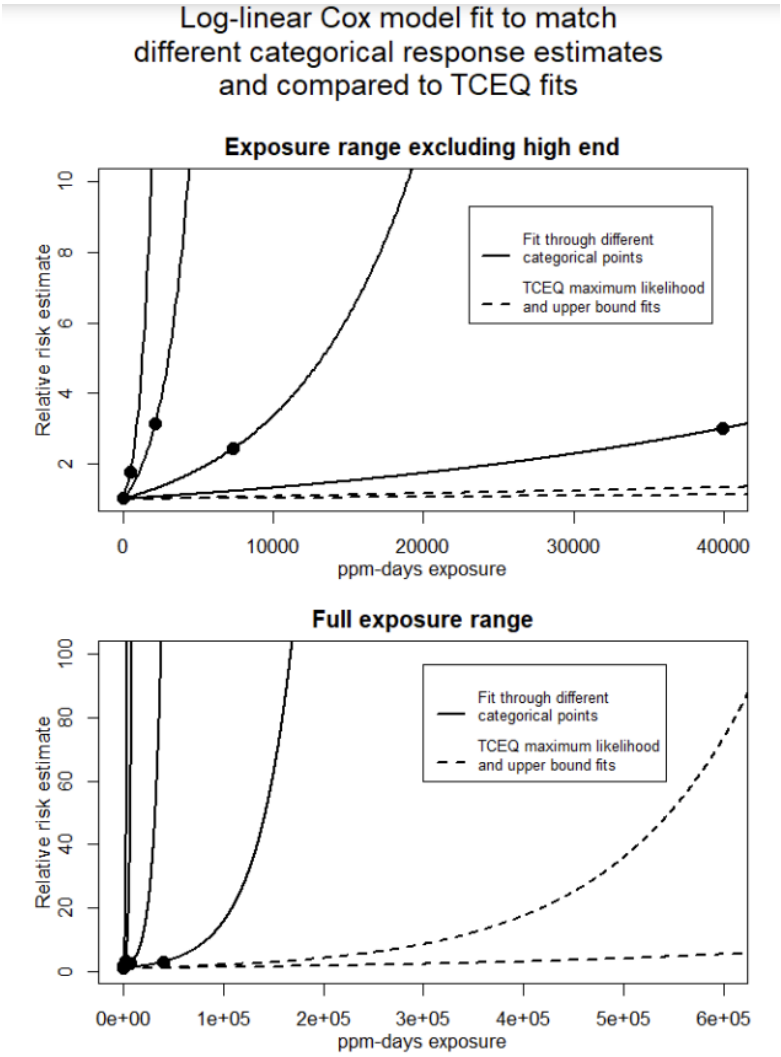
Table 1. Original and Corrected p-values for IRIS 2-slope linear spline and IRIS standard CPH model

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

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

EPA (2022) introduces a new visual evaluation of the lymphoid categorical data

EPA's (2022) response to comments provide new analysis of lymphoid data to attempt to prove TCEQ's standard CPH approach is a flawed approach. However, EPA makes several misstatements and errors in their analysis that (a) contradict EPA's own warning that comparisons of risk should not be made along the y-axis, (b) contradict the EPA SAB (2015) specific advice not to model categorical estimates as if they are reflective of individual data, and (c) equate biological plausibility with comparisons with the categorical estimates (EPA uses visual fit of the model to contend that the model is biologically plausible). This new analysis included the two following figures to prove that the CPH model "explodes upwards" at higher exposures. This analysis is incorrect for the following reasons:



1. EPA applies the CPH model to each categorical estimate using one categorical modeled estimate at a time. This approach is similar to taking data from animal studies and only modeling one mean value at a time to prove the model selected is incorrectly estimating the data at later exposure levels. This is irrelevant to the TCEQ analyses, because TCEQ did not analyze these data.

2. If EPA applied the same approach to the linear spline model it selected by focusing on the initial linear CPH model (below the knot), then this too would result in overprediction of the categorical estimates at later exposures.
3. This approach is diametrically opposite from the SAB's strong recommendation to NOT model the categorical data.
4. The approach continues to rely on visual fit of RR compared to the categorical estimates.
5. When the individual data are modeled as illustrated in the lower panel, the TCEQ model will begin to increase in estimating hazard rate and relative rate. However if we compare rate ratios using model equations, EPA's calculated RR is higher for the EPA's model than TCEQ's at higher cumulative exposures.

To address EPA's (2022) claim that the TCEQ model explodes upwards, the TCEQ RR model and EPA RR models were evaluated at intermediate, and the highest cumulative exposures in the NIOSH study to evaluate EPA (2022). The results show that if we use the TCEQ's and EPA's equations, the RR at 40,000 and 64,000 is higher for EPA than for TCEQ. In other words, if EPA (2022)'s claim that the TCEQ model explodes upwards is correct then it is even more so for EPA's model.

Table 2. Comparison of Rate Ratios Using TCEQ's and EPA's Models at higher cumulative exposures[1]

Exposure (ppm-days)	Rate Ratios	
	TCEQ	EPA
0	1.0	1.0
40000	1.12	2.60
642925	6.09	8.62

Graphical display of data is subject to manipulations not only in how data are expressed on the y-axis, but also with regard to the resolution of categorical models to represent the underlying individual data that were modeled. TCEQ (2020a) 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 exposure response. EPA correctly points out that each of these individual case categories will have very wide confidence intervals (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 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

[1] TCEQ DSD Table 6 RR = 1+exp(2.81E-6*ppm-days)
EPA IRIS (Table 4) RR = 1 + 7.58E-04*ppm-days - 7.48E-4*max(0, ppm-days-1600)
40000 ppm-days = max value plotted on Figure 4-3 of IRIS 2016
642925 ppm-days = max cumulative ppm-days in NIOSH study from Table 1 of Valdez-Flores et al 2010.

why comparison of the dose-response model results to the categorical model results is inappropriate on p. 52 of TCEQ's (2020b) response to the 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 by stating that the categorical model is a well-accepted method to represent the data. While it's true in epidemiology categorical approaches are used to understand patterns, it is also true that it depends highly on the number of categories. These data certainly should not be used as a basis for determining goodness of fit of the model to the underlying data, because the categorical estimates are NOT the underlying individual data. 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 numbers of categories.

TCEQ is not advocating the use of these graphs to assess visual fit as a method to select the models, but instead explaining why these graphs are misleading for determining fit of continuous models. Instead, TCEQ relies on more objective statistical modeling approaches to evaluate goodness of fit, rather than "eyeballing" comparisons using figures that are not fit for this purpose and that distort the true comparisons of models against the underlying individual data that were modeled.

EPA (2022) criticism of TCEQ graphs and explanations of visual comparisons along the y-axis

ACC (2020) supports TCEQ (2020) explanation of visual fit issues along the y-axis, which a TCEQ peer reviewer did not understand. Although it is apparent from the TCEQ peer reviewer and EPA's (2022) comments that the figures may have been confusing, the larger point is that they were meant to illustrate the warning note in EPA IRIS 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).*[2]*"* Part of the confusion may be that it is difficult to see the TCEQ (2020) label in the y-axis that change from Rate Ratio (or Relative Rate, RR) to Hazard Rates/Baseline Hazard Rate implied by the categorical estimates of RRs) when the standard CPH model was adjusted for differences in implied background hazard rates of categorical estimates of RRs and the linear two-piece spline to adjust for an approximation of the difference is based on line risks between the RRs and the CPH model.

Of greater concern is that EPA's (2022) response to TCEQ reveals a fundamental flaw in EPA's major reason for selecting the 2-spline model based on visual "local fit". EPA (2022; p. 53).

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

EPA (2022) states that the continuous modeling of the data involves comparison against the same reference group. EPA (2022) is justifying visual fit comparisons along the y-intercept based on an incorrect understanding that all models involve comparisons against the same reference group with the same estimated baseline hazard for all models. In other words, EPA (2022) appears to indicate that different models have the same implicitly estimated baseline risks because all the models compare the exposed group with the same reference group. This is incorrect and is in contradiction with EPA IRIS (2016) correct warning highlighted above that "*various models have different implicitly estimated baseline risks"*.

Proportional hazards modeling does NOT define a reference group against which data are compared. This is a major error that is at the root of all EPA (2022) statements regarding visual fit. The hazard rates that are modeled include individuals who have no estimated exposures (i.e. what EPA calls the "reference group") as well as individuals with different cumulative exposures.

Because each model is based on a specific equation, each model will have a different baseline background hazard rate (i.e. different implicit y-intercept). The y-intercept for each model depends not only on a "reference group" of individuals with no exposures, but also on the exposure response relationship (the equation for the model) and the individual responses at all exposure levels. This is why EPA IRIS correctly included warning footnotes to not use these graphs to compare along the y-axis. TCEQ attempted to graphically illustrate this in their figures because of public comments TCEQ received. TCEQ clearly states that the figures are an "approximation" of the y-intercept for based on the model equation and categorical models that better represent the 53 cases.

The SAS programs for Cox proportional hazards models do not estimate an intercept because the baseline hazard rate (the y-intercept for the hazard function) is factored out and not estimated. This is explained by a practical guide to survival analysis using SAS by Allison (2010[3]), which explains that "there is no intercept estimate . . .The intercept is part of $\alpha(t)$, the arbitrary function of time, which drops out of the estimating equations". In the graphs of relative rates (rate ratios), all the models are normalized to 1 at zero exposures by dividing the model's predicted hazard rate by the model's baseline hazard rate. The resulting rate ratios reported by SAS that are graphed in EPA IRIS figures are proportional changes to the baseline hazard rate **which is different for each model**.

In other words, it is impossible to make conclusions about over- or under- predicting the actual study results even if one were to incorrectly define the "actual data" as the 5 categorical estimates of RR. As explained above, this is because the baseline background hazard rates implied by the nonparametric (categorical) rate ratios and the underlying background hazard rates (HR) 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 as was done in the draft OEHHA (2023) and in public comments by the former

[3] Paul Allison, 2010: Survival Analysis Using SAS: A Practical Guide, second edition, P. 132, http://www.sthda.com/english/wiki/cox-proportional-hazards-model

chemical manager of the EPA IRIS (2016a) assessment. EPA's (2022) incorrect response to comments contribute further to this error.

Given EPA's (2022) misunderstanding, it becomes clearer why TCEQ (2020a) graphical explanation is indeed confusing to anyone who has an incorrect misconception that all continuous models involve comparisons against the same reference group with the same estimated baseline **hazard** for all models.

EPA (2022) cites a peer reviewer (anonymous expert #6) for the TCEQ draft assessment:

> "I don't understand why in Figure 22 the Cox regression was adjusted (by multiplying by RRo), thereby changing the estimate at zero dose to RRo, while the EPA spline model and EPA's categorical RRs were not adjusted and equaled 1.0 at zero dose. Based on the linear regression of the nonparameteric rate ratios for individual cases (as stated earlier I don't know how these were calculated), apparently, they were not relative to the unexposed category (because the RRo estimate is quite a bit greater than 1.0) as were EPA's categorical RRs, the Cox regression, and the EPA spline model. Therefore, I don't see the logic for adjusting only the Cox regression in Figure 22."

Part of the confusion could be that the TCEQ's draft assessment incorrectly labeled the y-axis as rate ratios (referred to by the peer reviewer as RRo). The TCEQ (2020a) corrected the y-axis from Rate Ratios to ratio of hazard rates (Figures 14 and 15 in the final TCEQ, 2020). Thus, it is inappropriate for EPA (2022) to use peer reviewer comments on a draft that was corrected. In addition, TCEQ's (2020a) focus is on comparing the standard CPH model with the categorical estimates of rate ratios to illustrate that it is incorrect to conclude based on the EPA IRIS (2016) figures that the standard CPH model is under estimating "the data".

In conclusion, TCEQ (2020) is not trying to defend the use of these visual fit figures as a means to select models. On the contrary, TCEQ (2020) is trying to illustrate why comparisons cannot be made along the y-axis, focusing primarily on comparing the CPH model with the categorical estimates of RR. While EPA is correct that categorical estimates of RR are often used to understand the pattern of the exposure-response, they are not the actual data modeled and should not be a basis for determining the fit of the model to the underlying individual data. Instead, TCEQ emphasizes a prediction analysis that are discussed in detail below. This TCEQ purpose is best expressed in 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."

2. **The epidemiological evidence does not support a steep exposure response. EPA dismisses all studies and findings based on external comparisons because of an**

HWE. EPA dismisses evidence from the UCC study because EPA incorrectly concludes the NIOSH study has superior exposure data and criticizes the UCC study for long follow up ignoring the fact that there were no findings at various follow up periods.

From an epidemiologic perspective, ACC commented on the implication from the IRIS model that EtO is a potent carcinogen. If real, a high risk at low concentrations should be easily detected in the most informative occupational epidemiological studies of exposed individuals, i.e. the NIOSH and UCC EtO studies. The effects of a potent carcinogen would be seen in both external and internal comparisons. Based on these two studies, there is no evidence that EtO is a highly potent carcinogen. EPA does not respond to the inconsistency between the implication of a potent carcinogen and the absent (UCC study) or weak associations (NIOSH studies) observed in external and internal comparisons. Furthermore, EPA dismisses any reliance on external comparisons (SMRs and SIRs) due to their concerns about the HWE. EPA also dismisses the lack of associations in the UCC and other studies because EPA (2016) and EPA (2022) considers the NIOSH study to be the only study with well-validated exposure analysis.

EPA (2022) incorrectly dismisses any reliance on external comparisons (SMRs and SIRs) due to the HWE.

EPA criticizes the Marsh et al. 2019 meta-analysis for relying solely on external comparisons and not considering internal from Steenland et al. 2003, 2004 (EPA MON Response to comments Dec. 2022 p.31). In fact, the authors did consider these analyses but discounted them for several reasons (Marsh et al. 2019; pp. 41-42 and 44-45), specifically, reliance on the log cumulative exposure metric, evident for only one gender (lymphoid tumors) and inconsistent with published results in Valdez-Flores et al. 2010. Dismissal of Marsh et al. is another example of EPA's reluctance to place any weight on external comparisons to the general population, even in the absence of the Healthy Worker Effect (HWE), as is the case for the NIOSH cohort of men and women (EPA MON Response to comments Dec. 2022 p.36-37). This is contrary to good epidemiologic practice, as evidence by Steenland et al., who conducted these analyses prior to the internal ones and clearly stated his results from the general population comparisons and that HWE was not an issue.

EPA uses the HWE to dismiss non-positive findings in external comparisons from the most informative epidemiology studies. In EPA's response to comments document, the Agency lists the citations from ACC regarding the evidence that a HWE diminishes over time, is of concern in studies with short follow up, and with extended follow up, is generally associated with non-cancer, not cancer causes. They even refer to comments that noted Steenland et al.'s dismissal of the HWE in the NIOSH study. This is what Steenland et al. was referring to below when he noted the change observed from the original study of the sterilizer workers 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 approach of considering the biological plausibility based on the epidemiological weight of evidence is consistent with IARC (1999) textbook on cancer epidemiology, which notes, "the

study results may be more convincing if a similar association were observed for a number of different comparison groups. For instance, with some occupational cohorts both an internal comparison group (people employed in the same factory but having a different job) and the experience of the general population (national and local rates) may be used."

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."

However, these topics are not addressed in EPAs response section. Instead of addressing these comments related to the HWE, EPA continues to discuss generalities about the HWE and its potential impact in studies that suffer from this source of error. They then ignore the well-known characteristics of the HWE and persist in attributing non-positive SMR and SIR results from the NIOSH study to the HWE. In other words the Agency never responds to ACC's citations that support that a HWE is unlikely to exist for the relevant disease end point here, cancer, especially with longer follow up times. Citation to a general proposition and ignoring information directed at the specific facts does not meet EPA's obligations to engage with the comments before it.

We note that the Agency does contradict themselves in citing SMRs and SIRs when they are positive (EPA MON Response to comments Dec. 2022 p.33). The agency should offer a balanced weight of the evidence assessment, considering results from both external and internal analyses and consider the inconsistent findings in their uncertainty analysis.

EPA (2022) incorrectly considers the NIOSH study to be the only study with well-validated exposure analysis. Consequently EPA (2022) ignores the negative UCC cohort study as important epidemiological evidence that should inform selection of the dose-response model.

The NIOSH and UCC study are the strongest epidemiological studies with individual estimates of exposures. The epidemiological weight-of-evidence should inform the selection of the exposure response model to apply to the NIOSH study, especially since there is no statistical difference from the null hypothesis for either the EPA IRIS (2016) 2-slope spline model and the standard CPH model. However, EPA downgrades the UCC and all other epidemiological studies because the NIOSH study has superior estimates of exposures. ACC (2020) and previous comments provided detailed information that the UCC study has arguably better exposure information prior to 1978.

EPA (2022) states: "A commenter claimed that the pre-1978 occupational exposures used in the EPA's dose-response model were underestimated by Hornung et al. (1994)." This and related comments are correct and supported by Bogen et al. publication. The Hornung regression model used by EPA to estimate pre-1978 historical exposures for which no data were available led to earlier exposure concentration estimates being lower later concentration estimates.

Hornung made incorrect assumptions for a key variable calendar year in the model. EPA (2016, 2022)explains that Hornung fit the measurement data from 1976 to 1985 and found that the effect of calendar year on exposure estimates was maximal between 1976 and about 1978-79 and

reduced exposure estimates after that. Thus, EPA IRIS (2016) and EPA (2022) explain that the calendar year effect in the exposure model was fixed at 1978 levels for years prior to 1978. They claim that fixing the effect of calendar year is consistent with available data and reasonable given that the increasing awareness of EtO in the late 1970s could explain calendar year effect decreasing

Even a cursory review of Hornung makes clear that his model is based on an assumption regarding the lack of changes in practices prior to 1978. There is no attempt in the paper to validate the conclusion or to evaluate if knowledge of potential health risks other than cancer may resulted in changes in practices. Indeed it is clear from ACGIH value changes in worker exposure concentrations (discussed below) that concerns about workers and EtO exposures predate the concerns about cancer risks in the late 1970s. The Hornung predicted early worker exposure pattern is inconsistent with historic worker exposure guidance ACGIH 1948, 1957), hygiene data collected in other industries (e.g., Von Grote et al. 2003, 2006), and published literature cited in Bogen et al. which show an evolution in sterilization operations over time and higher EtO residues degassing from sterilized materials in early periods of operations (e.g., Bruch 1961, 1972; Stetson et al. 1976: FDA 1978).

All the data support decreasing trends in EtO exposure concentrations over time from the earliest operations through periods during the establishment of industrial exposure guidelines and regulatory limits based on new toxicity information (ACGIH provided an exposure limit for EtO of 100 ppm in 1948 and 50 ppm in 1957). In other words, EPA's criticisms of the Bogen et al modeling because interviews of sterilizer workers were not systematically documented disregards all the evidence cited that show higher exposures in earlier years.

We note from the outset that even if EPA's criticism of lack of clarity of Bogen were accurate and that the uncertainties are too great to rely on, that criticism ignores the major point of these earlier comments --- EPA's exposure model is inconsistent with a significant body of literature and other facts demonstrating significant changes in practices that impact worker exposure. This calls into question EPA's stated basis for preferring NIOSH over UCC and refusing to integrate the UCC data into its analysis (or as a basis for recognizing that the steep slope of the 2-piece spline model is implausible).

EPA states that there is uncertainty in the Bogen model projections of industry-wide exposures of sterilizer operators because the earlier time periods were above the ACGIH health criteria at that time. However, there is much greater uncertainty in the Hornung model predictions of early worker exposure concentrations of approximately 10-30 ppm during the period when ACGIH established TLV limits of 100 ppm, and later 50 ppm, to guide the reduction in sterilization industry worker exposure. An increasing pattern of early sterilization worker EtO exposure is implausible based on available evidence.

EPA (2022) criticizes Bogen for insufficient documentation of methodology for the model and not reporting results of interviews while failing to acknowledge that EPA's basis for exposure data, the Hornung 1994 provided no data, text or figures for model predictions prior to 1976. Specifically, Hornung (1994) provided no supporting documentation for the key Hornung model

assumption that calendar year, a surrogate for improvements in work practices, be fixed at 1978 for the more than three decades of sterilizer operations prior to 1978 for which workers were being evaluated in the mortality study. This unsupportable assumption infers that there were no improvements in equipment and work practices between the 1930s when sterilization operations started and 1978 that would have changed sterilization worker exposures to EtO.

EPA (2022) states that fixing calendar year variable at 1978 levels "was both consistent with the available data for exposure levels prior to 1978 and reasonable given the increasing awareness of EtO carcinogenicity in the late 1970's …" Consistent with available data prior to 1978 means only data for 1976 and 1977, not the prior three decades and reasonable again only for the period for which data were available. In other words, Hornung provides no documentation for model predictions for worker exposures prior to 1976.

In contrast, Bogen et al. (2019) brought 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 operations 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.

EPA (2022) criticizes the Bogen estimates to be higher than ACGIH levels yet fails to explain why the NIOSH model would predict much lower exposures in the 1940's and 1950's. 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. One would not expect low exposure concentration when equipment and processes were crude, and little was known about EtO toxicity and no worker protection regulations. The NIOSH model predicted that the early worker increasing exposure pattern is inconsistent with industrial hygiene data collected in other industries (e.g., on Grote et al. 2003, 2006).

Assigning cases with underestimated exposures means that lymphoid cancer is associated with lower EtO levels than the workers with lymphoid cancers had been exposed to. Underestimating exposures associated with cancers will lead to an overestimation of potency. EPA (2016)

dismisses findings from the UCC and other studies because of the claim that the NIOSH exposure-response model is more reliable and better validated than other studies. The UCC exposure analyses has the advantage of exposures from another similar plant in earlier years. The evidence from both of these studies, individually and combined, do not support the 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 our ACC (2020, 2023) comments.

3. **Breast cancer, like other types of cancers 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 quantitative risk assessment based on substantial under ascertainment of incident cases reported by Steenland et al (2003) and subsequent risk deficits in the lower exposures.**

ACC raised the issue of under ascertainment of incident breast cancer cases in the Steenland et al. study (2003) and how a greater proportion of missing cases in the low exposure group of the interviewed population could result in an erroneous shape of the exposure-response curve (ACC MON comments 2020 pp33-35).

The issue of breast cancer incidence under ascertainment is one comment that was extensively but not adequately addressed by EPA. The Agency no longer just states case ascertainment is complete in the interviewed population as a response. Instead, they use the following arguments: 1) mortality analysis yields similar results but no under ascertainment, 2) similar results are seen in whole cohort and interviewed population and 3) the association of breast cancer with duration of employment further supports the exposure association (EPA MON Response to comments Dec. 2022; p. 28).

EPA's arguments center on what "supports the exposure association". Support for an exposure association tells you nothing about whether the data supports a 2-piece spline model, a CHP model or some other model all together.

We agree with the Agency that the NIOSH breast cancer findings of increased risk in the high exposure categories in both the mortality and incidence studies should be considered in cancer hazard weight of evidence. Indeed, ACC has not questioned inclusion in weight of the evidence. Rather our focus has been on the issue of missing cases as related to the use of questionable breast cancer incidence data for quantitative cancer risk assessment, i.e., to identify the shape of the exposure-response curve and calculate unit risk.

Our concern is differences in the low exposure groups between mortality data with no under ascertainment and incidence data that affects model choice. Comparisons of Table 5 of the mortality study and Table 3 of the incidence study show different patterns in the low exposure group. The later study shows a deficit (0.74) in the low exposure group (<647 ppm days), consistent with greater losses in low exposure group, while the mortality study does not

(SMR=1.0). It is not surprising that increases in the high exposure groups were seen in analyses of the entire cohort and the interviewed population.

Incident cases were missed in both the whole and interviewed subpopulations and are subject to the same potential bias of proportionally greater losses in the low exposure group. While it is true that duration of employment is correlated with cumulative exposure, the associations with breast cancer would not be expected to differ so dramatically, an observation that Steenland et al. 2003 pointed out. The results for the 15-year lag analyses in Table 5 show a coefficient of 0.039 and p value of 0.006 for duration versus 0.0000095 and 0.02 for cumulative exposure.

We are not arguing that an association has not been observed in the two studies and in the entire and interviewed population due to evidence of increased risk in the high exposure groups. Rather, we are arguing that there is also evidence of substantial missing cases that is proportionally more evident in low exposure (short duration employees) that erroneously suggests a steeper exposure-response curve. The breast cancer incidence data is so questionable that there cannot be sufficient confidence for its use in quantitative risk assessment.

EPA should reassess the issue of under ascertainment, as it relates to model selection, rather than arguing the presence of an association. Furthermore, if the agency feels compelled to include breast cancer as a target organ, despite the weak epidemiology evidence overall, they might consider using the NIOSH breast cancer mortality data for QRA that is fully ascertained and does yield a statistically significant increased risk in the highest exposure group with a 20 year lag.

EPA also uses Mikoczy (2011) to support supralinearity in the breast cancer incidence data. ACC (2020) submitted comments related to the Mikoczy study because it has been incorrectly cited as supportive of a supralinear association with breast cancer based on a comparison group that has an unusual deficit of the disease of interest (ACC MON comments 2020 p. 37). A referent group whose breast cancer rates are 50% of general population raises issues that can't be explained by the HWE.

EPA (2022) discusses the issues associated with the breast cancer study by Mikoczy et al. 2011 of Swedish female sterilant workers in the context of Marsh et al.'s discussion in their 2019 publication (EPA MON Response to comments Dec. 2022 p. 37-38). The Agency dismisses studies as relevant that suggest there is no HWE for breast cancer. Does that mean that EPA thinks there is a HWE for breast cancer and it is in the range of 50% for Mikoczy's control group of no or low EtO exposed working women? The agency doesn't clarify if this odd comparison group concerns them. Instead, they dismiss arguments of others who are concerned.

EPA claims it was being used to conclude that the referent group in the Mikoczy study should have the same baseline rates as the general population, which is incorrect. Fifty percent is more, however, than what could be explained by factors that may differ between workers and the general population. They state they "agree that definitive interpretation of the Mikoczy cohort results, taken by themselves may be challenging." Their discussion, however, does not clarify why they see this as challenging. We recommend that EPA discuss the specific limitations of

Mikoczy that limit interpretation and avoid using this study as a basis for supporting a steep exposure response model for breast cancer.

Breast cancer incidence is not an appropriate endpoint for quantitative risk assessment because the published data indicate a high potential for underascertainment of cases with shorter duration employment and, therefore, lower cumulative exposures. Furthermore, NIOSH does not allow access to breast cancer ***incidence*** data to allow independent analyses. ACC emphasizes that if human data are used, lymphoid mortality is the most appropriate endpoint for quantitative risk assessment because the weight-of-evidence for breast cancer is weak. However, if EPA includes breast cancer in the quantitative risk assessment, then breast cancer ***mortality*** is more appropriate because breast cancer mortality is fully ascertained (e.g. no missing data) and is publicly accessible. Based on EPA IRIS (EPA, 2016a, Table 4-11) application of the standard CPH model for breast cancer mortality, the central and lower-bound unit risk estimates for the cancer slope factor are 0.019[4] and 0.035[5] per ppm respectively, at the POD of 1/100[6] extra risk not including the EPA's age-dependent adjustment factor (ADAF). This is in contrast to the application of the 2-piece spline model for the under-ascertained breast cancer incidence of 0.72[7] and 1.5[8] per ppm for the central estimate and upper bound (see also EPA, 2016a, Table 4-15, not including ADAF).

[4] EPA IRIS (EPA, 2016a) Table 4-11 EC01(ppm) = 0.5305, not including ADAF

[5] EPA IRIS (EPA, 2016a) Table 4-11 LEC01 (ppm) = 0.285, not including ADAF

[6] Note: EPA IRIS provided no justification for applying the POD of 1/100 extra risk. An appropriate POD should be selected to ensure it is appropriately in the lower range of the experimental data.

[7] EPA IRIS (2016) Table 4-15 EC01(ppm) = 0.0138, not including ADAF

[8] EPA IRIS (2016 Table 4-15 LEC01(ppm) = 0.00675, not including ADAF

4. The biological evidence does not support a steep exposure response and should be the major driver for selection of exposure response models, rather than EPA's focus on log cumulative models.

As described in great detail in our previous comments, the epidemiological evidence and biological evidence are more consistent with the TCEQ model than the IRIS model. These comments build on earlier comments and more directly address recent EPA (2022) responses to public comments on the MON. Our earlier comments indicate that the EPA IRIS emphasis on statistical and visual fit is flawed and cannot be used as a basis for selecting between the TCEQ and IRIS model. Thus, biological plausibility is of paramount importance.

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. On one hand, EPA states that their evidence integration to inform the MOA for hazard assessment was adequate and appropriate. On the other hand, the agency argues that "not enough is known about the MOA to use evidence integration to inform the shape of an exposure response model representing cumulative lifetime human exposure." These contradictory positions precluded the integration of valuable information on biological plausibility into risk assessment. Instead, the IRIS exposure-response assessment is driven by comparisons of numerous statistical models without regard to biological plausibility. The EPA (2005) carcinogen risk assessment guidelines captures the general issue of applying multiple curve-fitting models based on statistical modeling approaches that applies to the EPA IRIS (2016a) assessment for EtO:

> "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 an exposure response form that is both biologically plausible and consistent with the observed data." Based on this statement it is clear that the SAB was looking for EPA to justify its model based on more than the statistical and visual fit approaches discussed above. 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 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 exposure 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 exposure 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 the shape of an exposure-response curve based on epidemiology data.

In fact, EPA (2022) argues that the exposure response information obtained from animal genotoxicity and carcinogenicity studies would not allow selection of exposure response models for human risk assessment for EtO. This position seems to be at odds with the agency's 2005 carcinogen risk assessment guidelines stating that "[i]f dose-response analysis of nontumor key events is more informative about the carcinogenic process for an agent, it can be used in lieu of, or in conjunction with, tumor incidence analysis for the overall dose-response assessment." Except for studies investigating sister chromatid exchanges (SCE), which are no longer considered as reflective of investigating genetic damage (Wilson and Thompson, 2007), the exposure response information for other genotoxicity endpoints from animal studies should be considered in the selection of the model for EtO risk assessment.

The putative genotoxic mode of action and toxicokinetic data should be a major consideration when selecting a model for risk assessment

While there is clear evidence that EtO is genotoxic and an animal carcinogen, 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 exposure 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 exposure 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 exposure response is not consistent with the biology of how EtO is hypothesized to work as a direct acting genotoxicant as evidenced from studies conducted in rodents. In their response to public comments, the EPA (2022) states that "EPA does not agree that observed exposure response pattern at high dose in rodent genotoxicity or carcinogenicity studies will generally be predictive of exposure response patterns in humans." While it is true that the exposure response observed at high doses in rodent studies often overestimate the hazard potential at low, environmentally relevant exposures, they seldom underestimate the hazard at the lower exposure levels for a vast majority of chemicals and most certainly for direct acting (i.e., those that do not require metabolic activation) DNA-reactive agents. Secondly, even in those so-called high dose studies, there are often lower dose levels with no detectable effects supporting the argument that the slope at the lower end of the exposure response cannot be steeper than at the higher doses.

The exposure response curves for EtO-induced gene mutations in the bone marrow (Recio et al., 2004, Figure 7) and lung (Manjanatha et al., 2017, Figure 8) tissues of transgenic Big Blue mice illustrate the above point. These two tissues are selected to illustrate the exposure response because they represent targets for EtO-induced tumors. In both cases, there is no evidence for a steeper initial slope at the low end . In fact, the mutant frequency at 25 and 50 ppm EtO was not statistically different from the concurrent negative control group in Recio et al. study whereas only the 200 ppm, but not 100 ppm, was significantly elevated from the control in the Manjanatha et al. study. and the initial shallow exposure response pattern is more consistent with the CPH model than the 2-piece spline.

Figure 1. Dose-Response for EtO-Induced *lac*I mutations in Mouse Bone Marrow (6 h/day; 5days/week; 48 weeks from Recio et al., 2004)

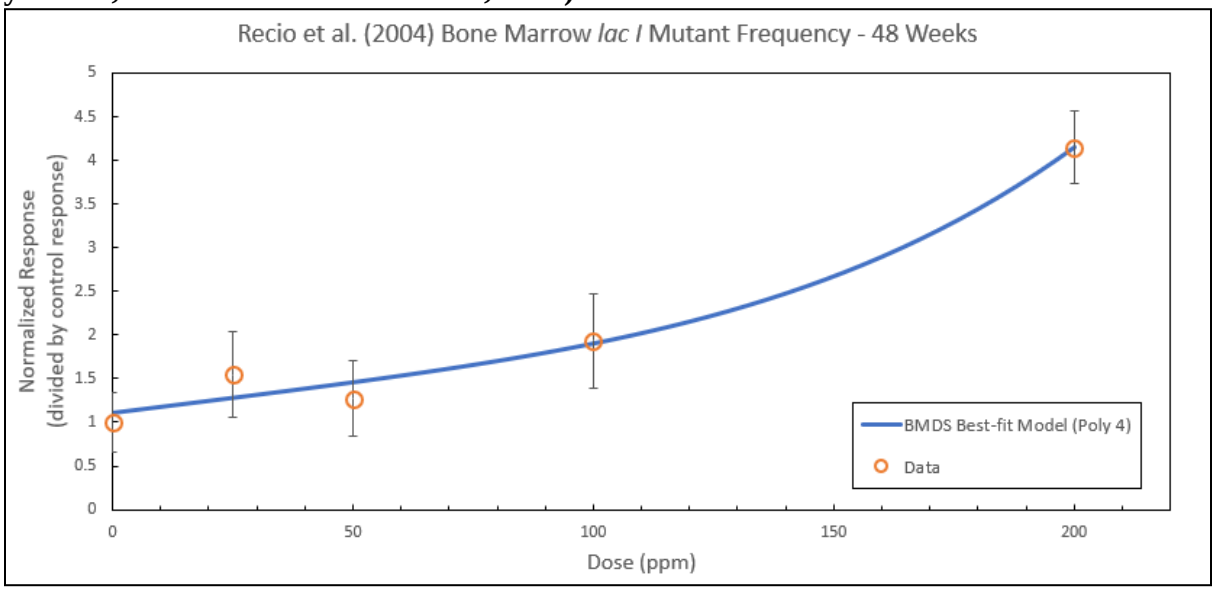
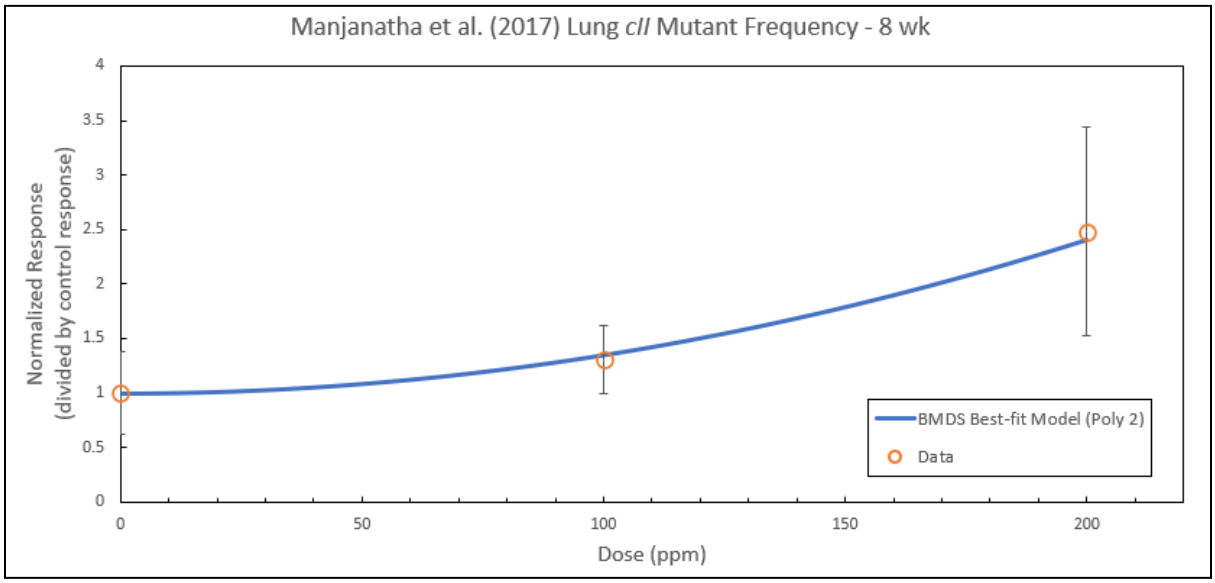
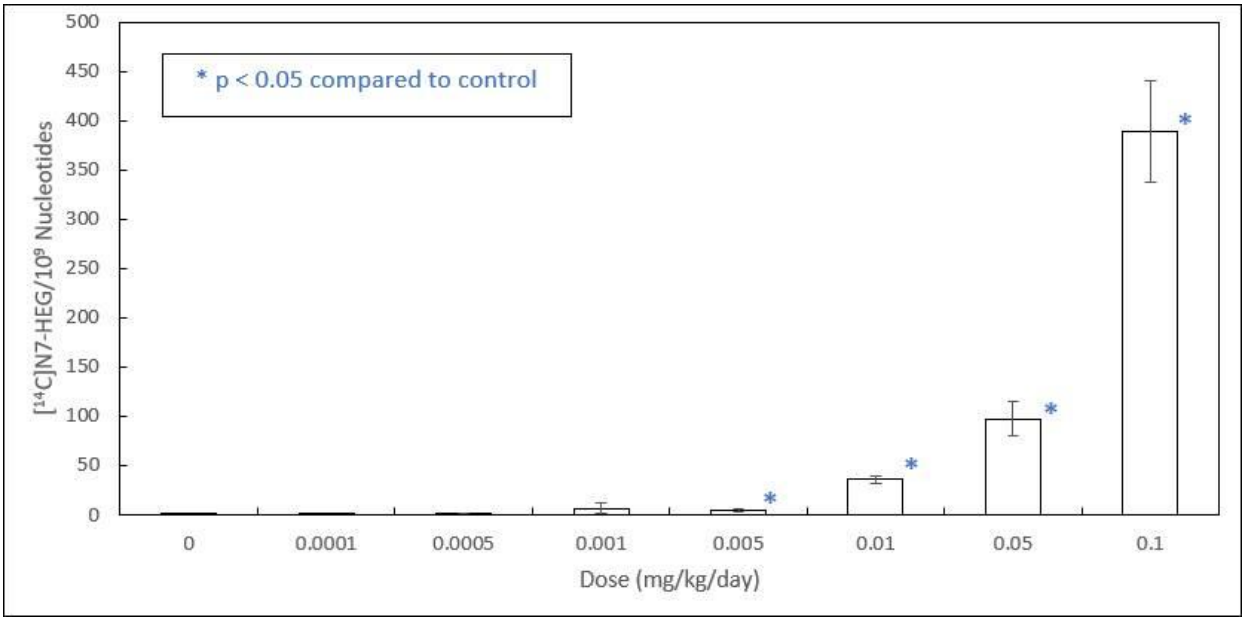


FIGURE 2. DOSE-RESPONSE FOR THE INDUCTION OF *cII* MUTATIONS IN THE LUNG TISSUE OF BIG BLUE B6C3F1 TRANSGENIC MICE AT 8 WEEKS OF INHALATION EXPOSURE TO ETO (MANJANATHA ET AL., 2017).



The above exposure 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). Marsden et al. have shown that significant increases in N7-(2-hydroxyethyl)guanine) (N7-HEG) adducts in the livers of rats treated i.p. with EtO were observed at the four higher i.p. doses, but not at the three lower doses (Fig. 9). This study using a wide range of dose levels provides no evidence for a steeper exposure response at the lower dose levels. The likelihood of such a steeper response at dose levels below the range investigated is virtually zero.

FIGURE 3. DOSE-RESPONSE FOR EXOGENOUSLY DERIVED DNA ADDUCTS IN LIVER OF [^{14}C]ETO-TREATED RATS MEASURED BY LC-MS/MS (MARSDEN ET AL., 2009).



EPA (2022) makes the following three statements in their response to public comments with regards to DNA adduct studies:

1. "…*much of the DNA adduct measurement studies evaluated by EPA were based on intermediate exposure durations (e.g., ≤4 weeks), which may not be informative of cumulative lifetime human exposures*..."
2. "… *Marsden et al. (2009) did not evaluate the effect of EtO exposure on DNA repair pathway activity, meaning the hypothesized repair at lower doses that is overwhelmed at higher doses was not confirmed*…"
3. "…*DNA adducts may not be the sole source of EtO-induced carcinogenesis*…"

EPA's concern that adducts data derived from studies with ≤4 weeks of EtO exposure duration might not be informative of cumulative lifetime human exposure seems to be speculative since no literature citation(s) were provided to substantiate their concern. Generally, the adducts are expected to reach a steady state, the duration of which is dependent upon the adduct type and the tissue in question. For example, the N7-HEG adduct has been shown to reach a steady state in 7-10 days whereas other adducts take a longer exposure duration (Pottenger et al., 2019).

With regards to the second statement, it is not clear why not having DNA repair data would change the fact that there was no significant increase in DNA adducts at lower doses in Marsden et al. study. Although DNA repair is one plausible explanation, other cellular protective mechanisms such as protein binding, GSH conjugation etc. also play a role in protecting DNA

from EtO-induced adducts. The most important take away from Marsden et al study is that the adducts levels induced at lower exposures of EtO are negligible and not what would have been predicted by the 2-piece spline model. Although the N7-HEG adducts measured by Marsden et al. are not considered mutagenic, they are the most abundant DNA adducts formed following EtO exposure (Walker et al., 1992). Thus, the shape of the exposure 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 O6-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).

EPA's statement that the adducts not being the sole source of EtO-induced carcinogenesis is accurate only in so far as the complex carcinogenic process is concerned. DNA adducts are biomarkers of exposure of the critical target and are not biomarkers of effect. However, for a direct acting alkylating agent such as EtO that is considered to act through a mutagenic MoA for its carcinogenic property, DNA adduction represents the molecular initiating event whose exposure response is informative of the subsequent events that follow depending upon the nature of the adduct, e.g., mutagenicity and carcinogenicity. Thus, the DNA adduct data together with genotoxicity data are informative, and EPA's rationale for dismissal of these data are at best speculative and inconsistent with the putative genotoxic mode of action.

In conclusion, the animal data support EtO as a weak genotoxic chemical and requires relatively high and prolonged exposures to induce mutagenicity. The experimentally observed exposure response patterns for mutagenicity/carcinogenicity show that the CPH model is biologically more plausible than the IRIS (EPA, 2016a) 2-slope model. Indeed, there is little support for a steeper initial slope in EtO exposure response for the key events (DNA adduction and genotoxicity) in the cancer MOA. 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.

5. **Both the CPH model and IRIS model estimate extra risk, so this fact should not be used as a basis to ignore reality checks based on valid estimates of endogenous levels. 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 at and substantially below that the body produces through natural processes.**

Endogenous production of EtO results from the oxidation of ethylene resulting from (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. 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).

EPA's (2022) response to public comments on the MON raised questions about the reliability of the Kirman et al. (2021) method of estimating endogenous levels. These comments favored speculative uncertainty and tangential information over the toxicokinetic and converging evidence from animals and humans that was presented in detail by Kirman et al. (2021) and by data in Filser and Klein (2019). As discussed in detail by ACC's response to the draft OEHHA IUR, endogenous levels can greatly exceed the levels to which residents near sterilization facilities may be exposed (ACC, 2023). While EPA's potency estimate technically only applies to exposures above endogenous levels, it is questionable that a chemical would be a potent carcinogen at levels that the body produces through natural processes.

We provide the following detailed responses to the EPA (2022) response to public comments (EPA comments are in italics)

p. 63: "*...The unit risk estimates the EPA developed are for extra risk (i.e., above background); background and endogenous levels of EtO, which would be relevant to (the true) background risk, are not integral to the development of the estimates of extra risk.*"

p. 64: "*IRIS risk estimate for EtO represents the increased cancer risk due to exposure to EtO emissions – above any potential existing risks from endogenous or ambient background levels of EtO exposure*"

- While this statement is technically true, it is misleading. In the interest of sound science, EPA has an obligation to assess and communicate total exposure and risk from EtO (including endogenous and background sources), as noted in CERLCA policy (EPA, 2002):

> *"In some cases, the same hazardous substance, pollutant, and contaminant associated with a release is also a background constituent. These constituents should be included in the risk assessment, particularly when their concentrations exceed risk-based*

concentrations. In cases where background levels are high or present health risks, this information may be important to the public. Background information is important to risk managers because the CERCLA program, generally, does not clean up to concentrations below natural or anthropogenic background levels."

Furthermore, this statement applies to both the TCEQ and EPA models because both estimate extra risk, but the TCEQ model is more reasonably conservative.

Page 64: "*however, all study participants were also exposed to ambient background levels of EtO in the air, meaning the "endogenous equivalent" value is an approximation of exposure to endogenous plus ambient background levels of EtO.*

- This criticism of Kirman and Hays (2017) is outdated and misleading. Adjustments for ambient exposures has been addressed in Kirman et al. (2021) in which the exogenous pathway for EtO in air was explicitly included in the estimates of endogenous exposures to EtO.

Page 64: "*The occupational exposures in the NIOSH study represent workplace EtO levels these workers experienced – and are in addition to any endogenous or broad population background exposures to which the workers may also have been exposed. Similarly, the relative risk estimates for workers are made in comparison to risks to individuals in the absence of occupational exposure – but these reference individuals would also have exposure to population background and endogenous EtO. Thus, methodologically, the levels of endogenous exposures (and baseline population level risks that may result from these exposures) have been accounted for in the EtO IRIS dose-response assessment.*"

- This statement is incomplete. While it may be reasonable to assume that endogenous exposures to EtO and ambient EtO exposures to workers in the NIOSH cohort are approximately the same as experienced by the current US population, this is not true for EtO exposures from smoking which are expected to be an important contributor to background EtO exposure. Specifically: 1) Smoking is an important source of EtO exposure (Kirman and Hays, 2017; NHANES data); 2) Smoking and associated EtO exposures were larger in the past (peaking in the 1950s-60s) than they are today; and 3) Smoking habits can differ between salaried and non-salaried workers (Hsu et al., 2019). For these reasons, smoking by NIOSH workers has the potential to confound EtO exposure estimates (rather than confounders on health response measurements). This effect is not accounted for in the analysis of the NIOSH cohort data and the application of potency estimates to the current US population.

Page 64-67: "*Additionally, the analysis was based on measurements of an exposure biomarker but did not include any direct measurements of endogenous EtO levels.*"

"ACC does not reference any direct measurements of endogenous EtO levels"

"*NHANES does not provide direct data on exposures to EtO.*"

- These statements are misleading. HEV measurements, as collected by NHANES, are well accepted as an appropriate and specific biomarker of EtO exposure (e.g., CDC refers HEV as "Ethylene Oxide Hemoglobin Adducts". As noted in EPA's IRIS document "*Formation of hemoglobin adducts has been used as a measure of exposure to EtO... Hemoglobin adducts are good general indicators of exposure because they are stable...*". HEV measurements, which are stable and reflect cumulative exposures to EtO, are a much better biomarker of EtO exposure than any direct measurements of EtO in blood. As such, "direct data" are not required to conclude with high confidence that exposure to EtO has occurred.
- EPA's ambient air monitoring data for EtO do provide "direct data" on exogenous exposures to EtO, and NHANES provides definitive data on total exposure to HEV.
- Indirect estimation of endogenous levels by subtracting exogenous contributions from total HEV burden (as performed in Kirman et al., 2021) is a scientifically sound approach.

Page 66: "*However, Kirman et al. (2021),72 which was funded by the ACC, as well as related publications by Lewis et al. (2022)73 and Sheehan et al. (2021),74 do not provide new information to substantiate the validity of proposed calculations of "endogenous equivalent air concentrations"*."

- This statement is factually incorrect. New data in Kirman et al. (2021), which significantly improve the assessment for calculating endogenous equivalents include the following:

    - NHANES HEV measurements made in >4,500 individuals (Table 1 of Kirman et al., 2021) are new data were not available in 2017. These new data replace pooled HEV data from control groups, and show that we underestimated total EtO exposures to the U.S. population in Kirman and Hays (2017)
    - EPA's measurements for EtO in ambient air (Table 6 of Kirman et al., 2021) are new data that were not available in 2017. These new data replace an assumption that the exogenous pathway is negligible and show that contributions to HEV are small (but not negligible). For this reason we underestimated contributions from this pathway in Kirman and Hays (2017).
    - Kirman et al. (2021) indicated that "*datasets are also available for characterizing the relationship between occupational exposures to ethylene oxide in air and HEV level (Boogaard, Rocchi, and van Sittert 1999; Brugnone et al. 1986; Duus et al. 1989; Hagmar et al. 1991; Norpoth and Bolt 1995; Törnqvist et al. 1986a), and are generally consistent with the correlation defined in Equation (1), within a factor of approximately 2.*" Although these data are not new, their use to justify the linear

relation between EtO exposure and HEV is a new application. Use of the Filser and Klein's (2018; Figure 5 below) along with the NHANES data as a function of cigarettes per day (see Figure 6 below) provide independent validation of the linear relationship used, and as such EPA should considered this relationship to be factually substantiated.

- In light of EPA's skeptical response on the endogenous pathway as characterized in Kirman and Hays (2017), we conducted an extensive review of the literature to characterize the underlying biochemical basis for HEV formation in preparing our follow-up paper (Kirman et al., 2021). These data help explain why HEV is readily detected in nearly all people and improves the biological plausibility of our assessment.

FIGURE 4. HEMOGLOBIN ADDUCTS (HEV, NMOL/G HB) IN WORKERS EXPOSED TO ETO IN AIR (FROM FILSER AND KLEIN, 2018).

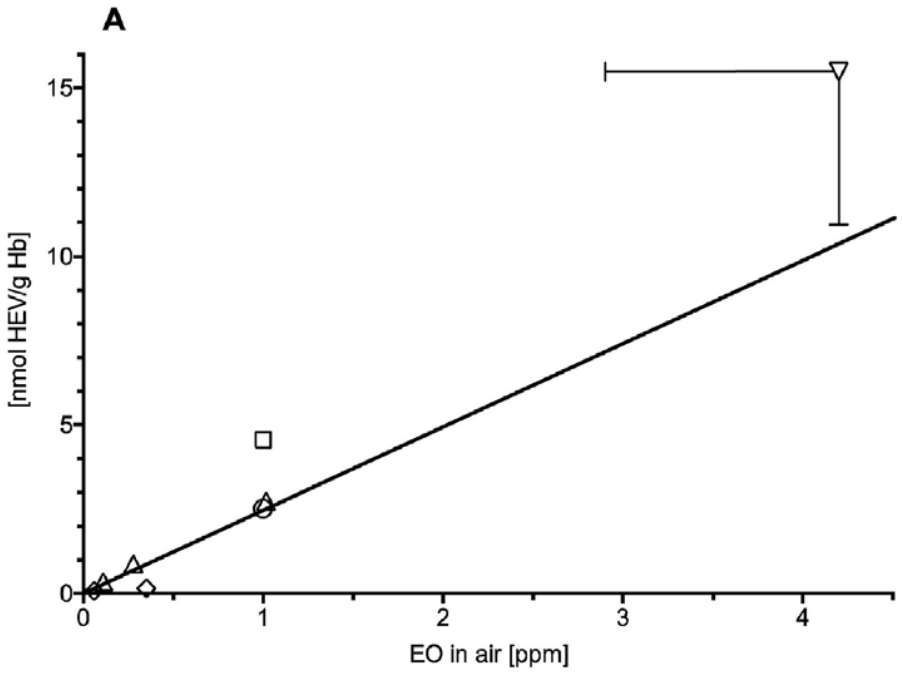
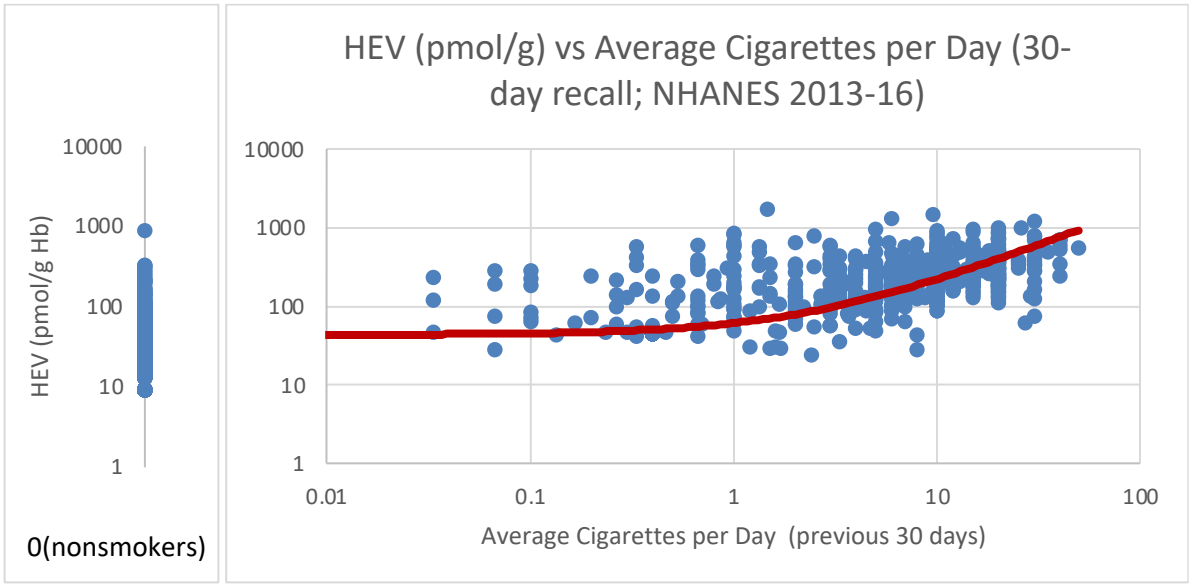


FIGURE 5. HEMOGLOBIN ADDUCTS (HEV, PMOL/G HB) IN U.S. SMOKERS EXPOSED TO ETO IN CIGARETTE SMOKE (NHANES, 2013-16).



p. 66: "*However, Kirman et al. (2021) utilizes this result to infer that a specific amount of human endogenous EtO can be attributed to production by gut bacteria. EPA considers this inference that EPA considers to be highly speculative*"

- The biochemical pathways described in Kirman et al. (2021) for endogenous production are expected to be operational in ALL mammalian species.
- HEV levels are detectable in mice, rat, and humans at levels that cannot be explained by exogenous exposures, which only account for a small fraction of the measured adduct burden. Therefore, endogenous exposures to EtO can be concluded with a high degree of confidence.
- Inferences from rodent studies are often made in human health risk assessment (e.g., absorption estimates based on rodent toxicokinetic studies; rodent tumors used to derive cancer potency estimates). There may be some uncertainty and variation with respect to the source of endogenous (GI lumen production vs systemic production), but this does not detract from the conclusions that both pathways are expected to be operable, and that endogenous production occurs in humans.

p. 66 "*a second investigation from the same laboratory using different dietary lipids (Kautiainen et al., 1991) provided little quantitative support for the hypothesis*"

- This statement is misleading. This study assessed a different oil (soya) compared to coconut oil. It did report a significant effect on HEV (when data were pooled together due to small sample sizes), albeit by a smaller amount than would be expected based on percent unsaturated fatty acid content of soya oil. The results of this study support the conclusion that dietary lipids can alter HEV burden.

p. 66: "*A more recent study (Paaradekooper et al, 2017) observed traces of ethylene exhaled by human volunteers administered an agent to induce an inflammatory response, which the authors attributed to peroxidation processes. However, in this experiment, both the baseline exhalation of ethylene and the increase attributed to the inflammatory agent were modest (roughly 0.5 ppb at baseline and a 0.5 ppb increase after treatment) - a finding that does not suggest a large source of endogenous exposure.*"

- This statement is misleading and incomplete. The low levels of endogenous ethylene production reported in this study are inconsistent with NHANES data, which should carry greater weight. In addition, there are other publications that address this topic that appear to report higher levels of ethylene in exhaled breath (e.g., Bratu, 2019; Cristescu et al. 2014; Popa et al., 2015). A more thorough review of available data on ethylene levels in exhaled breath is needed.

p.65-67: "*...this calculation amounts to a major and unvalidated change in the interpretation of the industrial hygiene relationship which specifically correlated increases in adduct levels above background with substantial occupational exposures to EtO.*"

"*however the applicability of the assumed relationship for "equivalent" air concentrations from Kirman and Hayes (2017) is simply assumed to apply to the NHANES adduct data.*"

"*Kirman et al (2021) also apply their hypothesized relationship between adduct levels and equivalent air concentrations*"

"*However, EPA concludes that these calculations represent a hypothesis as applied to population background adduct levels. As explained above, EPA does not believe that these assumptions have been factually substantiated.*"

- EPA's repeated use of qualifiers "unvalidated", "assumed", and "hypothesized" is a misleading characterization of the weight of evidence that supports our assessment. EPA appears to question two aspects of this assessment: (1) the applicability of biomarker data from occupational exposures (generally collected at high EtO concentrations) to the general population (considerably lower concentrations of EtO), and (2) the linear relationship observed between measured HEV and EtO in air. Use of occupational exposure data to make inferences for the general population is a common extrapolation (and ironically one that EPA has adopted to estimating cancer risk from EtO in the general population using data from occupationally exposed workers). In Kirman and Hays (2017), we adopted the relationship as characterized by Germany's DFG, which we considered a trusted, unbiased source to support our assessment.

- The linear correlation between HEV and EtO in air relied upon in Kirman and Hays (2017) and Kirman et al. (2021) is scientifically sound based upon multiple, independent sources of validation:

    - Physiologically based pharmacokinetic (PBPK) models are widely considered to be the "gold standard" for extrapolating from high doses to low doses in human health risk assessment. The best available PBPK model for EtO (Filser and Klein, 2018) predicts a linear relationship between EtO exposures in air and HEV burden (solid black line shown in Figure 5 above).
    - From an empirical standpoint, the HEV data from exposed workers, as summarized by Filser and Klein (2018; individual markers in Figure 5 above) and cited in Kirman et al. (2021), are also consistent with a linear relationship. The available worker data depicted in this figure indicate that a linear relationship between HEV adducts and EtO in air is at least maintained across a broad range of concentrations (~0.1 ppm to ~4 ppm).
    - 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 (Figure 6 above) depicted by redline; note - log-log scale which permits inspection of the data,

but also distorts the shape of a linear relationship at low exposures) between EtO exposure (using cigarettes per day as a metric) and HEV adduct formation:

- These data, as shown in Figure 6, indicate that the linear relationship between EtO exposure is maintained from background exposure levels up to 30x background levels. The linear relationship in this figure is also consistent with a linear relationship (e.g., assessed by multilinear regression analyses) for another EtO biomarker (urinary 2-hydroxyethyl mercapturic acid) as reported by CDC scientists (Kenwood et al., 2021). Together, these data provide strong and convincing evidence to support a linear relationship between HEV and EtO exposure as used in Kirman and Hays (2017) and Kirman et al. (2021).
- Because the exogenous exposures to ethylene 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 EtO, 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.

p. 68-9: "*As cigarette smoke contains many carcinogens, there is not a reason to expect, in advance, 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"*

- This statement is incorrect if you accept the EPA IRIS value. EPA's large unit risk value derived for EtO serves as the reason for expecting EtO contributes a large part of total lymphoid cancers risk due to smoking. EPA's unit risk value for EtO ($3x10^{-3}$ per ug/m^3) is **two to three orders of magnitude higher** than the unit risk values derived by EPA for any other lymphoid carcinogens present in tobacco smoke (e.g., benzene, 1,3-butadiene, formaldehyde).
- While accepting that smokers have EtO exposures and that according to their model, low exposures pose a steep breast and lymphoid cancer risk, logic concludes that the enormous number of studies that have examined cancer risks to smokers and the large number of studies of these specific diseases should have confirmed a causative relationship. None of the large number of carcinogens in cigarette smoke (about 70) are confirmed causes of lymphoid tumors or breast cancer based on epidemiology data, so they would not mask an EtO risk if it existed. Again, the high bar of quantitative analyses using statistical bounds is unreasonable and not necessary. Smoking is a common exposure in epidemiology studies and multiple non-positive studies in combination produce a narrow confidence interval. EPA failed to consider 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, who have been confirmed at high risk of numerous types of cancer, but not lymphoid tumors.

p. 70: "*We have found that canister age, materials used to line the inside of the canisters, and how the canisters are cleaned before they are put into service, can cause results to be biased high.*"

- If correct in a meaningful way, this statement only serves to increase the importance of endogenous exposure to EtO, for which EPA appears dismissive in other parts of the document. To the extent that EPA's air sampling data is biased high, our estimates of endogenous exposures (as presented Kirman et al., 2021), which utilize EPA's data to adjust for exogenous exposure contributions to HEV, would be underestimated. That is to say, endogenous production estimates may be biased low as a result of relying upon EPA's air sampling, thereby increasing the relative importance of the endogenous pathway to total EtO exposure.

p. 79: "*In the figures below, EPA provides graphs of observed dose-response shapes seen in animal inhalation cancer bioassays and in animal inhalation studies of mutagenesis and genotoxicity. For the cancer bioassay data, the graphs include a curve connecting responses for the lowest and highest dose response points. When intermediate data points line above this line it suggests a plateauing response. While when intermediate points fall below the line the graph suggests an upward curving response. If intermediate points fall on the line, the response is essentially linear. (Due to the property that these "quantile" tumor response data cannot exceed a probability of one, the basic "one-stage" linear hazard model is shown rather than a straight line.) The graphs for the mutation and genotoxicity studies can be similarly interpreted. (For these studies, a simple straight line is shown as do not have probabilities of effect approaching 1.) For both cancer bioassays and genotoxicity studies, these graphs include examples with concave-down (plateauing) shapes, examples with generally linear shapes, and examples showing upward curving patterns. The number of examples showing concave-down, plateauing, response shapes indicates that collectively these experimental cancer and mechanistic data are not consistent with claims from some commenters that such dose-response shapes are biologically implausible for EtO."*

- This statement is misleading. EPA is conflating dose-response plateaus at high concentrations in rodents with proposed plateaus at range of low concentrations (*e.g.*, a knot at 100 ppm-days can result from 100 days exposure to 1 ppm or from 10,000 days exposure to 0.01 ppm) that correspond to the knots assumed for cumulative exposure in their spline analysis. This conflation is not transparent due to different units used in rodents (ppm) and

epidemiology (ppm-days) data sets. There is no evidence of low-concentration plateaus in the animal data sets.

EPA (2022) suggested a "forward" analyses to validate the Kirman et al. (2021) model. ACC provides a new "forward" analyses that further validates external EtO exposures and internal EtO HEV hemoglobin adducts (EtOHEV)

The relationship between NHANES HEV biomonitoring data as a function of EtO exposure (using cigarettes per day) established by Kirman and Hays (2017) and Kirman (2021) is validated with a "forward" analysis, as suggested by EPA (2022), based on measured EtO concentrations in mainstream cigarette smoke. Using the linear relationship between external EtO exposures and internal EtO HEV hemoglobin adducts (EtOHEV), Kirman (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 EtOHEV 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 7 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 EtO 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)
Liu et al. 2014	3R4F	ISO	4.98
	3R4F	HCI	15.49
Forster et al. 2018	1R6F	HCI	10.23
	3R4F	HCI	11.48
Jaccard et al. 2019	1R6F	ISO	3.52
	1R6F	HCI	10.29
	3R4F	ISO	4.03
	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-EtOHEV endogenous-equivalent model approach.

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±4.3 ppb) along with the value of 17 CpD, results in an independently derived conversion factor of 0.53±0.25 ppb per CpD. Together these data indicate that the linear correlation between HEV in smoker exposures to EtO is in 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 EtO exposures. If the IUR is assumed as correct, a 10 ppb (10,000 ppt, as a representative midpoint from Table 7) external EtO 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), 5 orders of magnitude less than the approximate 10,000 ppt smoking exposures estimated by Kirman et al (2021) and validated by direct measurement of EtO in cigarettes. This suggests smoking should result in a readily demonstrable cancer signal, but the overall epidemiological data are weak or equivocal for this endpoint. 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.

6. TCEQ's prediction analyses provides an important reality check and is based on well accepted methods also used by IRIS (2016). This approach is superior to the subjective visual fit methods. Calculation of confidence intervals using a different method from TCEQ's results in nearly identical CI's and further support conclusions based on TCEQ's prediction analyses.

TCEQ's prediction analysis provides a superior, objective basis for selecting their model, compared to EPA's misleading visual presentation of data and incorrect statistics (described above). TCEQ essentially checks the two models by using the models and national mortality rates to estimate the number of cancer deaths that each model would predict in the NIOSH study.

Citing earlier comments, EPA rejects TECQ's analysis in part based on a claim that TCEQ only used the upper bound. In their final prediction analysis, however, TCEQ used both the maximum likelihood estimate and upper bound (p. 88) to demonstrate that TCEQ's model accurately predicts the cases.

EPA (2022) also dismisses the analyses because TCEQ used a method based on the inverse of the confidence intervals of the SMRs to estimate confidence intervals. Thus, it is useful to show that another well-accepted approach for estimating the 95% CI using the Exact Poisson method 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).

Table 4: Total NIOSH Cohort Lymphoid Cancer Mortalities Predicted by TCEQ (2020a) and EPA IRIS (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% Exact Poisson CI if the Model were True
TCEQ (CPH)	2.81E-06	52.42	98.9%	(40.1, 70.0)	(38.8, 68.2)
IRIS 2-slope spline 15-yr lag (MLE)	**7.58E-04**	**91.69**	**173.0%**	**(70.1, 122.4)**	**(73.3, 111.7)**

Note: There are 53 actual lymphoid mortalities. 53 is within the CI's for the TCEQ model but not within the CI's 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 Exact Poisson distributions (Ahlbom,1993). TCEQ (2020a, Table 6).

TCEQ (2020a) DSD[9] 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. The predicted number of cancers with confidence intervals based on the 95% Exact Poisson CI are shown in Table 5

Table 5: Total NIOSH Cohort Lymphoid Cancer Mortalities Predicted by TCEQ (2020a) and EPA IRIS (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% Exact Poisson CI if the Model were True
TCEQ (CPH)	2.81E-06	44.56	84.1%	(32.9, 59.1)
IRIS 2-slope spline 15-yr lag (MLE)	**7.58E-04**	**77.94**	**147.1%**	**(60.8, 96.2)**

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

[9] TCEQ Section A3.3.2

the actual number of cancers after including a theoretical HWE[10]. TCEQ used the inverse of the confidence intervals of the SMRs. We calculated the confidence intervals based on the Exact Poisson distributions (Ahlbom, 1993).

EPA (2022, p. 90) points to the SMR of 0.72 in Steenland et al. (2004) Table 3 in support an even larger potential HWE (i.e., 28%) in the NIOSH cohort for LH cancer deaths. This data is based on a very small sample size, 9 cases in males and 2 in females. This deficit is not statistically significant and is, therefore, consistent statistically with an SMR of 1.0.

It is important to note that TCEQ selected the 15% not because data or literature supports a HWE for cancer but because it was seeking to address EPA's concern that there might be a HWE (despite lack of data supporting one) To accommodate the concern, TCEQ selected 15-16% based on a Norwegian worker study with relatively short average follow-up of 11.5 yrs (Kirkeleit et al. 2013)[11]. This was done notwithstanding the fact that the study did not identify a HWE for breast or lymphoid cancer. Accordingly, the 15% HWE used in ACC's calculations is a very reasonable high estimate for any differences that might exist between the general US population and the NIOSH workers. We note again that this Kirkeleit study designed to address HWE found no HWE for lymphoid or breast cancer. Instead, TCEQ found the only statistically significant finding for all cancers as a basis for a possible 15%.

Although based on a study in Norway, ACC considers 15% to be reasonable for US population for cancers for the NIOSH study. The NIOSH study authors concluded that there was unlikely to be a cancer HWE in this longer follow-up study (Steenland et al. 2004[12])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 disease, being smaller for cancer than for other major diseases, and it tends to disappear with time since recruitment into the workforce (IARC, 1991). 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).

Based on a comparison of all causes of death among 10 cohorts of U.S. workers with the general U.S. population, Monson 1986 describes the initial phase of the HWE as the dynamic period, when the SMRs start out as deficits and increase over time until they hold in the stable plateau phase at 0.90, approximately 1.0 after 15 years. This is what Steenland et al. was referring to below when he noted the change observed from the original study published in 1991.

[10] Predicted is based on multiplying predicted values in Table 4 by 0.85 for HWE of 15%, and CI's calculated using Exact 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.

[11] 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.

[12] "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)

> "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."

Monson concludes his "Observations on the Healthy Worker Effect" as follows:

> "The healthy worker effect is relatively weak in comparison to causal excesses that can be detected in epidemiologic data."

Thus, 15% is a high estimate for HWE for lymphoid cancers and is a reasonable estimate for a sensitivity analyses to represent all possible differences between the NIOSH cohort and the general population. EPA (2022) insistence on using a non-statistically significant value based on limited sample from NIOSH study is unreasonable, especially in light of the peer-reviewed published conclusions by the authors that there is no HWE in the NIOSH cohort (Steenland et al. 2004).

In addition, a quintile analysis was also performed by TCEQ (2020a) to address EPA IRIS (2016) emphasis on the local fit of the models below the knot. 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 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 (Confidence Intervals Based on Exact Poisson Distributions.

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)[13] (7.7, 23.5)[14]	8.0 (4.5, 16.1) (3.5, 15.8)	9.4 (5.2, 18.8) (4.1, 17.1)	9.1 (5.1, 18.3) (4.1, 17.1)
Linear two-piece spline with knot @ 1,600 ppm-days – 15-yr lag (MLE)	**20.9** **(11.7, 42.0)** (12.2, 30.9)	17.6 (9.8, 35.2) (9.9, 27.2)	**20.8** **(11.6, 41.7)** **(12.2, 30.9)**	**20.9** **(11.7, 41.9)** **(12.2, 30.9)**

[13] TCEQ (2020a) used the inverse of the confidence intervals of the SMRs.

[14] For comparison, we calculated the CI based on Exact Poisson distribution(Ahlbom, 1993)

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. See TCEQ (2020a, Table 6).

DETAILED COMMENTS: PART 2

Risk management of EtO facility emissions is futile if managers do not take into count the background exposure that the U.S. population receives daily and apply risk management targets (i.e., Action Levels) that are, at a minimum, health-based and account for the total background exposure that the U.S. population receives daily.

- **Distinguishing facility EtO concentrations from background EtO concentration in ambient air is an important initial step in risk management and should follow EPA background guidance**

The first reality check in managing the risk of EtO from facility emissions is to determine statistically whether facility-related concentration exceed background concentrations at locations of interest around the facility. EPA has developed specific guidance for statistically comparing background and chemical concentration (EPA, 2002). Although this guidance was developed for chemicals in soils rather than chemicals in air, the approach, concepts and statistical tools are equally applicable to distinguish facility EtO concentrations from background concentrations. We recommend that EPA follow this guidance when assessing whether facility emissions are contributing significantly to ambient air beyond the fenceline.

- **Characterizing ambient EtO concentrations that constitute exogenous background exposure is critical to determining whether facility emissions contribute significantly to existing background EtO in ambient air.**

Characterizing background exposure concentrations is particularly important now as the Proposed HON endeavors to manage EtO residual based on the IRIS URE which provides essentially no risk management utility. Statistical measures of background concentrations provide a reality check of facility related concentrations in managing general population EtO risk. As EtO from natural and unregulated anthropogenic sources is found in ambient air throughout the U.S., everyone is exposed to EtO in ambient air regardless of where they live or work. Thus, characterizing and accounting for background EtO concentration is essential to distinguish industrial facility contribution to local air to inform risk management. It is critical that EPA properly account for background EtO concentrations in managing risk that is only associated with facility emissions.

A substantial amount of EtO concentration data has been collected by EPA, state and local agencies and emitting facilities to characterize background concentrations and in some cases facility-related concentrations. EPA has stated in response to previous comments that the "current method for measuring EtO cannot detect EtO at all levels" and in discussions they don't "trust" reported background EtO concentrations. The responses infer that there are reliability issues with EtO sample measurements. These criticisms come without documentation and are unsupported and misleading as monitoring for EtO and other regulated airborne chemicals often includes a percentage of samples with undetected concentrations and show temporal and spatial variability in concentration. However, these factors generally have no effect on reliable statistical characterization of concentrations. Indeed, a relatively small percentage of EtO

sample concentrations from the EPA NATTS/UAT monitoring program are reported as 0 (not detected), and neither do undetected sample concentrations nor sample variability limit the use of this monitoring data in reliably characterizing background EtO concentrations in ambient air.

There are considerable NATTS/UAT data to statistically characterize mean EtO concentrations and variability. Data from EPA's passive canister (original sampling method) and pressurized canister (refined sampling method) for recent years provide a reliably statistical characterization of current EtO ambient background concentrations. For years 2021 and 2022 there are NATTS/UAT monitoring data for EtO by the refined method from nine states and 23 individual monitoring locations (2607 samples). As 13 of 41 locations had monitoring locations reported >20% non-detected concentrations (reported as 0), ranges of mean and coefficient of variation (CV) estimates were calculated separately for the 41 locations (all data) and 28 locations with <20% ND (1888 samples) to assess the effects of higher percentage of ND. Mean concentrations for the entire data set ranged from 0.083-0.268 ppb and CVs from ~35-130%. Mean concentrations for the data set with fewer ND samples ranged from 0.097-0.268 ppb and (CV) from ~35-116%. Thus, the ranges were similar between data sets but slightly tighter for the subset with location reporting a higher percent of samples with detected concentrations as one would expect. Importantly, analysis of recent NATTS/UAT monitoring shows that mean EtO concentrations substantially exceed EPA IRIS risk specific concentrations (RSC; 0.0001-0.01 ppb) reaffirming the need to account for ambient background concentrations in managing risk for populations residing near emitting facilities.

EtO monitoring data from a background location (South Dekalb) for emitting facilities in the state of Georgia (part of NATTS/UAT data) has been used to statistically compare mean background concentrations with concentrations from near facility monitoring locations. We note that statistical comparisons of mean EtO concentrations for locations outside the fenceline of emitting facilities in Georgia has shown that concentration at these monitoring locations are, in many cases, indistinguishable from background concentrations (Georgia EPD, 2022; Sheehan et al., 2021; Lewis et al., 2022). Analysis of ambient background concentrations along with modeled facility concentrations for near residential locations also have also been used to evaluate the health significance of facility contributions to near facility population exposure (e.g., Lewis et al., 2022). These analyses confirm the importance of properly characterizing and statistically comparing background and facility contributed EtO concentrations in ambient air in informing management of general population EtO exposure and risk.

- **Characterizing total background exposure from combined exogenous and endogenous contributions is considered as an ultimate reality check in managing EtO risk from facility emissions**

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). Therefore, characterizing both endogenous and exogenous is critical to accounting for general population total exposure. 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. In the previous section, we addressed EPA (2022) recent comments on the Kirman et al. (2021) approach and provide a new "forward" analyses recommended by EPA (2022) to further validate the Kirman et al. (2021) approach. A table of selected percentiles of total and endogenous equivalent concentrations can be found in supplemental material with Lewis et al. (2022).

There are published examples of where total equivalent concentration comparisons have been useful in informing risk management decisions. Sheehan et al. (2021) compared 50^{th} and 90^{th} percentile concentrations from monitoring around eight facilities plus 50^{th} percentile endogenous equivalent concentration (total exposure concentrations) with 50^{th} and 95^{th} percentile total equivalent background concentrations for nonsmokers in the U.S. population to assess the relevance of facility contributions to total exposure. Similarly, Lewis et al. (2022) compared modeled concentrations at near facility residences to state of Georgia mean and median background concentrations and total exposure concentration based on modeled 5-year average EtO concentration at the near residence plus 50^{th} percentile endogenous equivalent concentration plus 50^{th} percentile background concentration for Georgia for eight emitting facilities in Georgia (total exposure) with the 50^{th} 60^{th} and 95^{th} total equivalent concentrations for the nonsmoking U.S. population. The comparisons showed that monitored or modeled concentrations, in nearly all cases, contributed negligibly to total background exposure. Both comparisons provide another level reality check on the significance of the facility emission contribution to near facility EtO exposure. A total background comparison would be a useful reality check

- **A representative Action Level for EtO would be health-based and would provide adequate health protection for populations residing near emitting facilities yet also inform whether additional risk management in required**

The EPA guidance for comparing background concentrations described above specifies that the difference from background mean must be "sufficiently large to warrant additional interest based on health or ecological information" (i.e., an Action Level should be related to health or ecological risk in a meaningful way; EPA, 2002).

Any health-based Action Level also must be based on considerations accounting for total background exposure (endogenous and exogenous) and its associated U.S. non-smoking population variability. Such considerations have been described by Kirman et al. (2021), who noted that the 50^{th} percentile background exposure to endogenously generated EtO is equivalent to a 2.3 ppb and that exogenous EtO exposure contributes ~0.2 ppb for a total equivalent EtO exposure concentration of 2.5 ppb. The 95^{th} percentile upper-bound total equivalent exposure

concentration is 5.5 ppb. Importantly, it was also determined that the endogenous-equivalent exposure constituted approximately 95% of background exposures, with the remaining 5% attributed to EtO present in background ambient air. Given these data, any consideration of a biologically defensible Action Level defining potential health risks of facility contributed EtO exposures must, at a minimum, be sufficient to be reasonably biologically differentiated from naturally produced endogenous EtO and its associated population variability.

In developing an Action Level for EtO emissions, an initial reality check should be an evaluation as to whether when considering the 50^{th} percentile background equivalent endogenous concentration (2.3 ppb), the added facility-related concentration is sufficient to exceed the upper bound 95^{th} percentile (5.5 ppb) total equivalent background EtO exposure for the U.S. nonsmoking population. This provides important exposure context as to whether the facility contribution is sufficiently large to exceed total background exposure. Accounting for a 50^{th} endogenous equivalent background concentration of 2.3 ppb and ~ 0.2 ppb exogenous ambient background concentration, suggests that a minimum Action Level for EtO would be ~3 ppb, which is this facility-related concentration when added to combined background would slightly exceed the upper bound total equivalent background concentration. Although we do not believe that this minimum health-based Action Level for EtO is the best health risk-based value, it does appropriately account for total background exposure and provides starting point for discussion of a health-risk based value.

Considering the lack of risk management utility provided by the IRIS RSCs, an alternative health-risk-based assessment model is needed to support a health risk-based Action Level for EtO consistent with EPA background guidance. Fortunately, the biologically plausible and peer reviewed TCEQ risk assessment for EtO, which was developed subsequent to the EPA risk assessment, is based on the same NIOSH data, and considers total exposure, provides a reasonable alternative basis for an Action Level. TCEQ RSCs for 1 in 1,000,000 to 1 in 10,000 risk range from 0.24 to 24 ppb. Understanding that the size of local populations residing near emitting facilities is relatively small, basing an Action Level on some fraction of the 1 in 10,000 RSC of 24 ppb, would be reasonable and scientifically valid approach to setting a health risk-based Action Level for EtO emissions. Instead, a TCEQ health risk-based Action Level would be consistent with EPA background guidance and should be protective of potentially exposed populations but sufficiently large so as to warrant consideration of further action to reduce facility emissions. Such an Action Level would provide practical utility in managing residual risk of facility EtO emissions

EtO **References**

Ahlbom A. 1993. Biostatistics for Epidemiologist. CRC Press. Boca Raton. p.56. https://books.google.com/books?id=cEwPEAAAQBAJ&pg=PA56&dq=poisson+exact+confidence+interval&hl=en&newbks=1&newbks_redir=1&sa=X&ved=2ahUKEwip1YPrzfX_AhUMElkFHcQ7D9kQ6AF6BAgJEAI

American Chemistry Council (ACC). 2020. Comments on the proposed amendments to the "National Emission Standards for Hazardous Air Pollutants: Miscellaneous Organic Chemical Manufacturing Source Category" (MON). American Chemistry Council Ethylene Oxide Panel. March 19. https://www.regulations.gov/comment/EPA-HQ-OAR-2018-0746-0151

American Chemistry Council (ACC). 2023. ACC Comments on Draft Cancer Inhalation Unit Risk Factor (IUR) for Ethylene Oxide. June 14. Available at: https://oehha.ca.gov/media/dockets/21069/21111-acc_eo_panel/ca-_final_oehha_eo_iur_comments_jun_2023.pdf

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.

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

Boogaard PJ, Rocchi PS, van Sittert NJ. 1999. Biomonitoring of exposure to ethylene oxide and propylene oxide by determination of hemoglobin adducts: Correlations between airborne exposure and adduct levels. Int Arch Occup Environ Health 72: 142-150.

Bratu AM. 2019. Spectroscopic study of breath ethylene via the mouth and nose. Lasers Med Sci 34(4): 773-778.

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.

Brugnone F, Perbellini L, Faccini GB, Pasini F, Bartolucci GB, DeRosa E. 1986. Ethylene oxide exposure. Biological monitoring by analysis of alveolar air and blood. Int Arch Occup Environ Health 58: 105-112.

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.

Cristescu SM, Kiss R, Hekkert St, Dalby M, Harren FJ, Risby TH, Marczin N; Harefield BIOSTRESS study investigators. 2014. Real-time monitoring of endogenous lipid peroxidation by exhaled ethylene in patients undergoing cardiac surgery. Am J Physiol Lung Cell Mol Physiol 307(7): L509-L515.

Duus U, Osterman-Golkar S, Törnqvist M, Mowrer J, Holm S, Ehrenberg L. 1989. Studies of determinants of tissue dose and cancer risk from ethylene oxide exposure. In: Proceedings of the Symposium on Management of Risk from Genotoxic Substances in the Environment, 141-153. Solna.

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.

Forster M, Fiebelkorn S, Yurteri C, Mariner D, Liu C, Wright C, McAdam K, Murphy J, Proctor C. 2018. Assessment of novel tobacco heating product THP1.0. Part 3: Comprehensive chemical characterisation of harmful and potentially harmful aerosol emissions. Regul Toxicol Pharmacol 93:14-33.

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.

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

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.

Hagmar L, Welinder H, Lindén K, Attewell R, Osterman-Golkar S, Törnqvist M. 1991. An epidemiological study of cancer risk among workers exposed to ethylene oxide using hemoglobin adducts to validate environmental exposure assessments. Int Arch Occup Environ Health 63: 271-277.

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.

International Agency for Research on Cancer (IARC). 1999. Cancer Epidemiology: Principles and Methods. 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.

Kautiainen A, Törnqvist M, Anderstam B, Vaca CE. 1991. In vivo hemoglobin dosimetry of malonaldehyde and ethene in mice after induction of lipid peroxidation. Effects of membrane lipid fatty acid composition. Carcinogenesis. 1991 Jun;12(6):1097-102.

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.

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.

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.

Liu X, Joza P, Masters A, Rickert B. 2014. Determination of Ethylene Oxide (ETO) in Mainstream Cigarette Smoke Using Hydrobromic Acid Derivatization and Gas Chromatography-Mass Spectrometry Method. Presented at 2014 CORESTA Congress, October 12-16, 2014, Québec City, Canada.

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.

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.

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

National Research Council (NRC). 2007. *Models in Environmental Regulatory Decision Making.* Washington, DC: The National Academies Press.

Norpoth K, Bolt HM. 1995. Ethylene oxide. In: The MAK Collection Part IV, BAT value documentations, 2, 146-161. Wiley VCH, Weinheim, Germany.

Paardekooper LM, van den Bogaart G, Kox M, Dingjan I, Neerincx AH, Bendix MB, Beest MT, Harren FJM, Risby T, Pickkers P, Marczin N, Cristescu SM. 2017. Ethylene, an early marker of systemic inflammation in humans. Sci Rep 7(1): 6889.

Popa C, Petrus M, Bratu AM. 2015. Ammonia and ethylene biomarkers in the respiration of people with schizophrenia using photoacoustic spectroscopy. J Biomed Opt 20(5): 57006.

Pottenger LH, Boysen G, Brown K, Cadet J, Fuchs RP, Johnson GE, Swenberg JA. 2019. Understanding the importance of low-molecular weight (ethylene oxide- and propylene oxide-induced) DNA adducts and mutations in risk assessment: Insights from 15 years of research and collaborative discussions. Environ Mol Mutagen. 60(2):100-121.

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.

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.

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

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, Seals R, Klein M, Jinot J, Kahn HD. 2011. Risk estimation with epidemiologic data when response attenuates at high-exposure levels. Environ Health Perspect 119(6):831-7.

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

Swenberg JA, Fryar-Tita E, Jeong YC, Boysen G, Starr T, Walker VE, Albertini RJ. 2008. Biomarkers in toxicology and risk assessment: informing critical dose-response relationships. Chem Res Toxicol 21(1): 253-65.

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

Törnqvist M, Mowrer J, Jensen S, Ehrenberg L. 1986a. Monitoring of environmental cancer initiators through hemoglobin adducts by a modified Edman degradation method. Anal Biochem 154:255-266.

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. Environmental Protection Agency (EPA). 2002. Guidance for Comparing Background and Chemical Concentrations in Soil for CERCLA Sites. Office of Emergency and Remedial Response U.S. Environmental Protection Agency. Washington, DC 20460. EPA 540-R-01-003, OSWER 9285.7–41.

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). 2019. IRIS EtO Assessment – Modeling Comparisons and Assessment of Uncertainty. Memorandum from Kristina A. Thayer, Director, Chemical & Pollutant Assessment Division (CPAD), ORD Center for Public Health and Environmental Assessment (CPHEA) to Peter Tsirigotis, Director, Office of Air Quality Planning and Standards (OAQPS), Office of Air & Radiation (OAR); dated October 18, 2019. U.S. Environmental Protection Agency, Office of Research and Development, Research Triangle Park, NC. Available at: https://www.epa.gov/sites/production/files/2019-11/documents/memo_sensitivity_of_ethylene_oxide_risk_estimates_to_dose-response_model_selection_c_.pdf

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 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, Ethlyene 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, 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.

Valdez-Flores, C. and R. L. Sielken. 2013. Misinterpretation of categorical rate ratios and inappropriate exp*osure–response model fitting can lead to biased estimates of risk: Ethylene oxide case study. Regulatory Toxicology and Pharmacology, 67, pp. 206-214.

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 Grote J, Hürlimann C, Scheringer M, Hungerbühler 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.

Wilson DM, Thompson LH. 2007. Molecular mechanisms of sister chromatid exchange. Mutat Res 616:11–23.

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