Ethylene Oxide
Carcinogenic Dose-Response Assessment

CAS Registry Number: 75-21-8

Prepared by
Joseph T. Haney, Jr., M.S.
Allison Jenkins, M.P.H.
Jessica L. Myers, PhD
Toxicology, Risk Assessment, and Research Division

Development Support Document

Revised Draft, January 31, 2020

TEXAS COMMISSION ON ENVIRONMENTAL QUALITY

DSD History

Effective Date	Reason
August 16, 2017	Public request for toxicity information
June 28, 2019	DSD proposed for public comment
January 31, 2020	Revised DSD and responses to public comments on 2019 draft DSD posted
To be determined, 2020	DSD posted as final

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Acronyms and Abbreviations

Acronyms and Abbreviations	Definitions
AMCV	air monitoring comparison value
°C	degrees Celsius
DNA	deoxyribonucleic acid
DSD	development support document
EC	effective concentration
ESL	effects screening level
$^{chronic}ESL_{nonthreshold(c)}$	chronic health-based effects screening level for nonthreshold dose response cancer effect
EtO	ethylene oxide
HEV	hemoglobin N-(2-hydroxyethyl)-valine
IARC	International Agency for Research on Cancer
LCL	lower confidence limit
LEC	lower limit on the effective concentration
LHN	lymphohematopoietic neoplasms
Lymphoid cancer	Includes leukemia (and specifically myeloid and lymphocytic leukemia), non-Hodgkin's lymphoma, and multiple myeloma
MW	molecular weight
µg	microgram
$\mu g/m^3$	micrograms per cubic meter
mg	milligrams
mg/m^3	milligrams per cubic meter
MLE	maximum likelihood estimate
mm Hg	millimeters of mercury
MOA	mode of action
n	number
N/A	Not applicable
NATA	National Air Toxics Assessment

Acronyms and Abbreviations	Definitions
NEI	National Emissions Inventory
NIOSH	National Institute for Occupational Safety and Health
NHL	non-Hodgkin's lymphoma
OSHA	Occupational Safety and Health Administration
PEL	permissible exposure level
POD	point of departure
ppb	parts per billion
ppm	parts per million
ppt	parts per trillion
RR	risk ratio
SAB	Science Advisory Board
SAS	Statistical Analysis System
SCE	sister chromatid exchange
SD	standard deviation
SE	standard error
SIR	standardized incidence ratio
SMR	standardized mortality ratio
TCEQ	Texas Commission on Environmental Quality
TD	Toxicology Division
UCC	Union Carbide Corporation
URF	unit risk factor
USEPA	United States Environmental Protection Agency
USFDA	US Food & Drug Administration
WHO	World Health Organization
WV	West Virginia

Chapter 1 Key Findings, Executive Summary, and Summary Tables

Executive Summary

- Ethylene oxide (EtO) is a chemical with many industrial applications and is particularly useful as a sterilant for medical devices.
- Because EtO is emitted in Texas and has been determined to be a carcinogen, the TCEQ undertook a carcinogenic dose-response assessment and derivation of a unit risk factor (URF) and an effect screening level (ESL) for this chemical.
- The TCEQ dose-response assessment considers new data and/or analyses from the scientific literature not available in 2016 (e.g., Vincent et al. 2019, Marsh et al. 2019, IARC 2019, Kirman and Hays 2017) as well as new TCEQ analyses and new data provided to TCEQ (e.g., dose-response model predictions of the underlying lymphoid cancer data, evaluation of the potential for healthy worker effects for EtO-specific cancer endpoints, sensitivity analysis of model predictions of the underlying cancer data to healthy worker effects for overall cancer mortality, as of yet unpublished summary results from a recent cohort update, Cox proportional hazards modeling results for multiple exposure lag durations, comparison of endogenous doses to occupational carcinogenic doses and environmental risk-based doses).
- Review of the EtO literature demonstrated that EtO operates by a direct-acting mutagenic mode of action (MOA) and suggests that the EtO cancer dose-response should be no more than linear overall with sublinearity expected by both the TCEQ and USEPA (2016) at endogenous levels and below.
- In addition, EtO is produced endogenously and an ambient air concentration of ≈1.3 ppb would be required to increase the internal dose of EtO by 1 standard deviation. Therefore, ambient EtO concentrations significantly less than 1 ppb (e.g., USEPA's acceptable air concentrations of 0.0001-0.01 ppb) would not be expected to produce biologically meaningful internal doses considering the range of normal endogenously-produced background EtO levels.
- Consistent with TCEQ guidelines (TCEQ 2015), recently derived toxicity factors and guideline air levels were reviewed to determine if there is a toxicity factor or guideline air level that is suitable for adoption by the TCEQ. As such, the USEPA's recently completed Evaluation of the Inhalation Carcinogenicity of Ethylene Oxide (USEPA 2016) was reviewed. The USEPA derived a URF of 9.1E-3 per ppb (lymphoid and breast cancer, ADAF adjusted), which corresponds to a 1 in 100,000 excess cancer risk air concentration of 0.001 ppb.
- The human data available for deriving an EtO toxicity factor are from two very high exposure occupational cohorts (Union Carbide Corporation (UCC) and National Institute

for Occupational Safety and Health (NIOSH)) that provide no information about the shape of the dose-response curve at low (i.e., environmentally-relevant) EtO concentrations. The TCEQ agrees with USEPA's determination that in the low-dose (i.e., endogenous) range a sublinear dose-response is "highly plausible," based on the MOA and information about endogenous production of EtO.

- In contrast to their determination that the low-dose (i.e., endogenous) region of the EtO dose-response curve is highly plausibly sublinear, USEPA ultimately chose to model EtO-induced lymphoid cancer with an overall supra-linear two-piece spline model that has a very steep linear slope in the low-dose region.
- The TCEQ evaluated USEPA's URF and overall supra-linear (i.e., linear two-piece spline) modeling choice in the context of the available observed data to determine the validity of the modeling and URF:
    - Endogenous Levels of EtO – USEPA's URF estimates that ambient concentrations of EtO > 0.01 ppb would produce an unacceptable increased cancer risk of greater than 1 in 10,000. This estimated ambient EtO concentration corresponds to an internal dose that is over 30 times lower than the 1st percentile of normal endogenous background levels (non-smokers), which is highly unlikely to be biologically meaningful and is inconsistent with the assessment of excess risk.
    - Population-Level Lymphoid Cancer Risk – Using measured concentrations of a biomarker of internal EtO exposure (an EtO-specific protein adduct in blood), it can be estimated that the mean amounts of background EtO levels would be equivalent to ambient concentrations of EtO of 1.9 ppb in non-smokers and 18.8 ppb in smokers. Accordingly, at measured internal background levels of EtO, the USEPA's URF for lymphoid cancer (4.8E-03 per ppb, ADAF unadjusted; incorporating age-dependent adjustment factors (ADAFs) in the exposure scenario) would predict a population-wide lymphoid cancer incidence rate greater than the USEPA-cited lymphoid cancer background incidence rate of 3% (in the absence of any other potential causes of lymphoid cancer). Because the population-wide lymphoid cancer incidence rate would have many contributing factors, not just a single chemical, this indicates that USEPA's selected model assessment overestimates observable lymphoid cancer risk based on endogenous/background levels of EtO alone.
    - Lymphoid Cancer Risk from Cohort Studies – The UCC cohort shows no statistically significant increased risk of lymphoid cancer with EtO exposure. The NIOSH cohort shows statistically significant increased risk of lymphoid cancer mortality at relatively high cumulative exposures. These data are not consistent with USEPA's selected model assessment (i.e., upper bound on the linear two-piece spline model) because that model assessment would predict statistically increased risks at even the lowest EtO cumulative exposures (see below).

- Model Fit with Observed Data – USEPA conducted their EtO cancer dose-response modeling using the NIOSH cohort data. To verify that USEPA's final selected model assessment (i.e., upper bound on the linear two-piece spline model) properly fit the original data, it was used to predict the expected number of lymphoid cancer deaths based on the same NIOSH individual exposure data as USEPA used for modeling. *Whereas 53 lymphoid cancer deaths were observed* in this cohort of 17,530 workers, *USEPA's selected dose-response model assessment predicted 141* (95% confidence interval (CI) of 108, 188) lymphoid cancer deaths in this same cohort. A sensitivity analysis assuming a healthy worker effect for overall cancer mortality (despite cancer endpoint-specific data to the contrary) also found that USEPA's model (maximum likelihood estimate) statistically significantly over-predicted the total number of lymphoid cancers for the cohort. Similarly, USEPA's final selected model assessment statistically significantly over-predicts lymphoid cancer deaths in every cumulative exposure quintile and indicates that statistically increased lymphoid cancer mortality should have occurred in every exposure quintile (including the lowest), when in fact this did not occur. This demonstrates unequivocally that USEPA's selected model assessment cannot be validated by the data that was used to derive it, and this model is not appropriate to use for estimates of population risk.

- The TCEQ determined that USEPA's use of an overall supra-linear dose-response model to derive their URF: 1) is not justified by the MOA data (which support a no-more-than linear dose-response); 2) is not consistent with predicted population risk from endogenous EtO for lymphoid cancer; and 3) statistically significantly over-estimates the number of lymphoid cancer deaths in the cohort from which the dose-response model was derived. **Therefore, the TCEQ found that USEPA's EtO inhalation URF is not adequately supported by scientific data (consistent with Vincent et al. 2019) and the TCEQ did not adopt it for this evaluation.**
- The TCEQ conducted a systematic review for studies that could inform the derivation of a cancer URF for inhalation exposures to EtO. This review identified key epidemiological data from two cohorts of occupationally-exposed workers, and Cox proportional hazards modeling was conducted to model the EtO-cancer dose-response.
- The TCEQ ultimately chose lymphoid cancer mortality as the critical cancer endpoint, using a 15-year EtO exposure lag with results for NIOSH males being more conservative than males and females combined, to calculate an **ADAF-unadjusted URF of 2.5E-6 per ppb (1.4E-6 per µg/m^3)** and an ADAF-unadjusted air concentration of **4.0 ppb (7.1 µg/m^3)** at an excess cancer risk level of 1 in 100,000 (TCEQ 2015).
- The elimination of breast cancer as a key cancer endpoint for EtO is consistent with, for example: (1) Recent studies evaluating the strength of the overall weight of evidence for

EtO-induced breast cancer (Vincent et al. 2019, Marsh et al. 2019); (2) A recent IARC (2019) analysis evaluating tumor site concordance across species that found "the state of the science does not support tumour site concordance as a general principle"; i.e., laboratory animal data are not relevant for supporting specific cancer sites in humans, particularly for breast cancer (as was done by USEPA 2016) based on results reported in IARC (2019); and (3) TCEQ's consideration of normal endogenous EtO doses and the biological plausibility/implausibility of modeled risk-based doses being associated with excess risk.

- As with USEPA's URF, the TCEQ's URF was evaluated in the context of the available observed data to determine the validity of the modeling and URF:
    - Endogenous Levels of EtO – Compared to endogenous EtO levels, the TCEQ's ADAF-unadjusted 1 in 100,000 excess risk air concentration of 4.0 ppb would produce an internal exposure equivalent to between the 90^{th}-95^{th} percentile of the normal endogenous range and could biologically plausibly be associated with excess risk above and distinguishable from normal endogenous EtO contributions to background risk.
    - Population-Level Lymphoid Cancer Risk - At measured internal background levels of EtO, the TCEQ's URF would predict a population-wide lymphoid cancer rate that is well within the background population lymphoid cancer rate (unlike USEPA's URF).
    - Lymphoid Cancer Risk from Cohort Studies – The standard Cox proportional hazards model of lymphoid cancer mortality did not show a relationship with EtO exposure that was statistically significantly different from zero. Therefore, by assuming a significant positive slope in the EtO-cancer association, the TCEQ is making a conservative decision to assume that EtO is causing lymphoid cancer in the exposed workers in the NIOSH cohort. Adding to this conservatism is the TCEQ's decision to use an upper confidence limit on the slope.
    - Model Fit with Observed Data – To verify that the TCEQ's model properly fit the original data, the expected number of lymphoid cancer deaths based on the individual exposure estimates for the NIOSH cohort (also used by USEPA) were calculated. Whereas *53 lymphoid cancer deaths were observed* in this cohort of 17,530 workers, *the TCEQ's selected dose-response assessment (i.e., upper bound of the Cox proportional hazard model) predicted 59* (95% CI of 45, 78) lymphoid cancer deaths. A sensitivity analysis assuming a healthy worker effect for overall cancer mortality (despite cancer endpoint-specific data to the contrary) also found that TCEQ's model neither statistically significantly over- or under-predicted the total number of lymphoid cancers for the cohort. Similarly, TCEQ's selected assessment neither significantly over- or under-estimated lymphoid cancer deaths for any exposure quintile. This demonstrates that the TCEQ's

model selection provides a superior fit to the observed number of lymphoid cancer deaths in the NIOSH cohort.

- The application of ADAFs resulted in an **ADAF-adjusted URF of 4.1E-06 per ppb (2.3E-06 per µg/m^3)** and an **ADAF-adjusted $^{chronic}ESL_{nonthreshold(c)}$ of 2.4 ppb (4.3 µg/m^3)** at an excess cancer risk level of 1 in 100,000, which **would produce an internal exposure approximately equivalent to the 75th percentile of the normal endogenous range.**
- The TCEQ determined that the use of Cox proportional hazards models to derive a URF for inhalation EtO cancer risk: 1) is justified by the MOA data showing EtO to be a direct-acting carcinogen whose effects, particularly at doses near the endogenous range, would be buffered by cellular repair mechanisms; 2) is consistent with population background risk considering background internal EtO levels (i.e., does not overestimate population risk for lymphoid cancer mortality); and 3) accurately estimates the number of lymphoid cancer deaths in the cohort from which the dose-response model was derived. **Therefore, the TCEQ's ADAF-adjusted URF for EtO has a sound scientific basis and will be adopted for review of air concentration data and for use in air permit reviews.**

Summary of Key Points

In 2016, the USEPA derived an inhalation URF for EtO (9.1E-03 per ppb or 5.0E-03 per µg/m^3; p. 4-91 of USEPA 2016) based on an overall supra-linear two-piece spline model (USEPA 2016). The URF is primarily driven by USEPA's dose-response assessment of lymphoid cancer in the NIOSH cohort. Despite extensive review by the USEPA Science Advisory Board (SAB) and extensive public comments on the science, in this Development Support Document (DSD) the TCEQ is able to demonstrate that *the model assessment ultimately selected by USEPA* (i.e., the upper bound of the linear two-piece spline model with the "knot" at 1,600 ppm × days, 15-year exposure lag) *statistically significantly over-estimates the actual number of lymphoid cancer mortalities observed in the NIOSH cohort; predicting 141 (95% CI of 108, 188) if USEPA's model were accurate versus the 53 actually observed (Figure 8). By contrast, the model assessment selected by the TCEQ* (i.e., the upper bound of the Cox proportional hazards model, 15-year exposure lag) based on relevant considerations discussed herein *predicts 59 lymphoid cancer mortalities versus the 53 actually observed*. Furthermore, *the USEPA's model assessment statistically significantly over-predicts lymphoid cancer mortality in every NIOSH cumulative exposure group, whereas the TCEQ's model assessment neither statistically over- or under-predicts for any cumulative exposure group (Figure 9 through Figure 12).*

Supra-linear models are generally not biologically plausible and tend to grossly overestimate low-dose risks. Therefore, sufficient mechanistic or biological data are required to support the application of a supra-linear model (i.e., the steep lower-dose component) for low-dose extrapolation (TCEQ 2015). USEPA (2016) provides no solid mechanistic or biological foundation

for adopting an overall supra-linear dose-response model, particularly its steep slope in the range of interest (e.g., typical environmental levels). In fact, *USEPA acknowledges the lack of mechanistic data to support the biological plausibility of a supra-linear dose-response,* stating *"the EPA is not aware of a mechanistic explanation"* and citing *"insufficient information to elucidate a basis"* (pp. I-29 and I-34 of USEPA 2016). Indeed, all the relevant considerations (e.g., MOA, normal endogenous background levels) discussed in various sections of this DSD consistently support the conclusion that there is a lack of data to adequately support the application of a supra-linear model with its steep low-dose slope to extrapolate to significantly lower (e.g., ambient air) EtO doses. Moreover, the available dose-response data from the NIOSH cohort (e.g., Steenland et al. 2004) are not informative as to the shape of the dose-response curve across doses of true interest (e.g., in the range of typical environmental concentrations), which USEPA acknowledged (p. I-14 of USEPA 2016; see Section 3.4.1.4.1). Workers were exposed to extraordinarily high concentrations of EtO, with exposure means ≈1,000,000-2,000,000 times higher than central tendency environmental exposures (animal carcinogenicity data are at even higher mean concentrations) and daily job exposures ranging from ≈15,000-32,000,000 times higher than central tendency environmental exposures. High-dose carcinogenicity data alone are incapable of informing truly low-dose risk, no matter how extensive the analyses or peer review (i.e., other relevant information such as mechanism/MOA must be duly considered). Indeed, USEPA acknowledges that *"the actual exposure-response relationship at low exposure levels is unknown"* (pp. 4-61 and 4-74 of USEPA 2016). *USEPA (2016) should not have based a URF on a supra-linear model (i.e., its lower-dose component) without a robust mechanistic justification for expecting its steep low-dose slope at truly low doses nor should the USEPA have used it to make a large low-dose extrapolation across an area (i.e., the endogenous range) where the agency in fact considers sublinearity "highly plausible."*

However, USEPA did ultimately derive a URF based on a supra-linear model (i.e., the lower-dose slope of the linear two-piece spline model), which necessarily leads to the following implausible conclusions when considering endogenous levels of EtO:

- *The air concentration at the maximum acceptable excess risk (0.01 ppb at 1E-04 risk) corresponds to an internal dose that is over 30 times lower than even the 1st percentile of normal endogenous background levels* (see Section 3.4.1.2.2.2);
- Expressed in other terms, based on USEPA (2016) and data on normal endogenous background levels, *air concentrations corresponding to more than ≈0.5% percent of mean normal endogenous background levels in non-smokers would be considered to be associated with unacceptable risk*; and
- *The predicted lymphoid cancer incidence based on mean background levels alone is greater than the population background rate cited by USEPA* (see Section 3.4.1.2.1.1).

The statistically significant overestimation of risk, *driven by a lymphoid cancer model for which there is inadequate statistical evidence that the slope is even greater than zero in the NIOSH (or UCC) cohort* (Appendix 4), undermines accurate risk communication and can lead to unintended societal consequences (e.g., medical supply shortages per the USFDA).

Consistent with the discussion above, the TCEQ has derived an inhalation URF for EtO because currently available information indicates that the existing USEPA URF results in biologically implausible risk estimates at environmentally-relevant air concentrations where use of the steep low-dose slope from an overall supra-linear two-piece spline model is not justified. For example:

- *The air concentrations corresponding to the USEPA acceptable excess risk range* (1 in a million to 1 in 10,000 based on the USEPA 2016 URF) *are orders of magnitude below those corresponding to the normal endogenous background range* (see Figure 7), making minuscule contributions to internal exposure that are not biologically meaningful as resulting total exposures are indistinguishable from normal endogenous background;
- Statistically significant increases in critical cancer endpoints observed in the NIOSH cohort occur at carcinogenic cumulative exposures that are *orders of magnitude above* endogenous levels (below which USEPA extrapolates orders of magnitude), and USEPA had no truly low-dose data to inform the shape/slope of the dose-response over the normal endogenous background range much less near typical environmental or risk-based air concentrations, which are even lower;
- The *biological implausibility of an overall supra-linear model for extrapolating risk down to endogenous levels* (and lower environmental and risk-based levels) is in fact supported by USEPA stating, "EPA considers it *highly plausible* that the dose-response relationship over the endogenous range is *sublinear*"; contrary to USEPA's expectation, the USEPA (2016) URF is based on a supra-linear model (i.e., the steep lower-dose linear component) that was used to extrapolate over the endogenous range and even below; and
- Consequently, *when USEPA's URF is used in conjunction with population-weighted mean background levels in non-smokers and smokers, a background lymphoid cancer incidence greater than the USEPA-cited background incidence is predicted based on background EtO levels alone* (see Section 3.4.1.2.1), which suggests a *scientifically unreasonable* URF**.**

Moreover, USEPA may not have adequately explored the potential contributions of ethylene to EtO risk, stating "only ≈3% of exogenous ethylene was converted to EtO in workers exposed to 0.3 ppm" and that "exogenous ethylene exposure is unlikely to contribute significantly to the effects associated with exposure to exogenous EtO in humans" (p. 3-30 of USEPA 2016).

However, based on USEPA's URF, mean environmental concentrations of ethylene in many areas would in fact result in unacceptable excess risk estimates when multiplied by 0.03 to account for the USEPA-cited endogenous conversion of 3% of exogenous ethylene to EtO. More specifically, greater than a 1E-04 excess risk would be estimated by USEPA's EtO URF at long-term ethylene air concentrations greater than 0.37 ppb. Interestingly, mean ethylene concentrations reported in human breath (e.g., 23 ppb in Fenske and Paulson 1999, baseline mean of 29-32 ppb reported in Bratu 2019) exceed this 1E-04 excess risk concentration by over 60-fold. These and other considerations are discussed in more detail within this DSD. *The TCEQ concludes that available information (e.g., mechanistic, biological) does not adequately support use of a supra-linear model (i.e., the steep lower-dose slope of the linear two-piece spline model) for extrapolation to truly low (e.g., environmental) air concentrations*.[a] Consequently, the TCEQ conducted a systematic review of the EtO human cancer literature and derived an appropriate URF consistent with the available data (e.g., MOA, model performance predicting the underlying cohort cancer data, model fit criteria, endogenous data).

The TCEQ ADAF-unadjusted URF for EtO based on lymphoid cancer is 2.5E-06 per ppb (1.4E-06 per µg/m^3) and results in an ADAF-unadjusted risk-based air concentration of 4.0 ppb at the no significant excess risk level of 1 in 100,000 (TCEQ 2015). The internal dose from continuous exposure to this EtO air concentration would correspond to the upper end of the endogenous range (i.e., between the 90th and 95th percentile), which is more biologically plausibly consistent with the assessment of excess (i.e., above and distinguishable from EtO background) risk. *The TCEQ-selected dose-response assessment (i.e., upper bound of the Cox proportional hazards model, 15-year exposure lag) accurately predicts the underlying cohort dose-response data (Figures 8-12)*. **The application of ADAFs resulted in an ADAF-adjusted URF of 4.1E-06 per ppb (2.3E-06 per µg/m^3) and an ADAF-adjusted $^{chronic}ESL_{nonthreshold(c)}$ of 2.4 ppb (4.3 µg/m^3)**.

Table 1 provides a summary of the risk-based value from a chronic, carcinogenic evaluation of EtO for use in air permitting and air monitoring. Please refer to Section 1.6.2 of the *TCEQ Guidelines to Develop Toxicity Factors* (TCEQ 2015) for an explanation of the various values used for review of ambient air monitoring data and air permitting. Table 2 provides summary information and the physical/chemical data of EtO.

---

[a] The TCEQ's finding that USEPA's EtO inhalation URF is not adequately supported by scientific data is consistent with the recent study Vincent et al. (2019), who conclude that "the USEPA's derivation of the IUR for EtO using a 2-piece, supralinear dose-response model—giving rise to one of the highest cancer potency estimates—appears not to be adequately justified based on the published literature and deviates from USEPA standard risk assessment guidance" and that "the IUR derived by USEPA grossly overestimates risk."

Table 1: Chronic Health-Based Screening Values for EtO

Screening Level Type	Duration	Value 1 (µg/m^3)	Value 2 (ppb)	Usage	Flags	Surrogated/ RPF	Critical Effect(s)	Notes
$^{chronic}ESL_{nonthreshold(c)}$ [a]	**70 yr**	**4.3**	**2.4**	**P,M,R**	**A,S,D**	**--**	**Lymphoid cancer in occupationally-exposed workers**	**--**

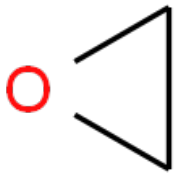
Bold values used for air permit reviews; values have been rounded to two significant digits.

[a] Based on the ADAF-adjusted URF of 4.1E-06 per ppb or 2.3E-06 per µg/m^3 and a no significant risk level of 1 in 100,000 excess cancer risk.

Usage:
P = Used in Air Permitting
M = Used to Evaluate Air Monitoring Data
R = Used to Calculate Remediation Cleanup Levels
N = Usage Not Defined

Flags:
A = AMCV report
S = ESL Summary Report
D = ESL Detail Report

Table 2: Chemical and Physical Properties

Parameter	Value	Reference
Molecular Formula	C_2H_4O	ATSDR 1990
Chemical Structure		ChemSpider 2019
CAS Registry Number	75-21-8	ATSDR 1990
Molecular Weight	44.05	ATSDR 1990
Physical State at 25°C	Gas	ATSDR 1990
Color/Form	Colorless gas	ATSDR 1990
Odor	Sweet, olefinic	ATSDR 1990
Synonyms	Ethylene oxide; oxirane; epoxyethane	ATSDR 1990
Solubility in water	$1x10^6$ mg/L	ATSDR 1990
Log K_{ow}	-0.22	ATSDR 1990
Vapor Pressure	$1.095x10^3$ mmHg	ATSDR 1990
Melting Point	-111°C	ATSDR 1990
Boiling Point	11°C	ATSDR 1990
Conversion Factors	1 ppm = 1.83 mg/m^3 1 mg/m^3 = 0.55 ppm	ATSDR 1990

Chapter 2 Major Sources and Uses

EtO is used as a chemical intermediate in the manufacture of ethylene glycol (antifreeze), polyester, detergents, polyurethane foam, solvents, medicine, adhesives, and other products. The conversion of EtO to ethylene glycols represents a major use for ethylene glycol in the US (IARC 2012). Relatively small amounts of EtO are used in sterilization of surgical equipment and plastic, as a fumigant, and as a sterilant for food (spices) and cosmetics (IARC 2012). In 2018, EtO was being produced in the US at 15 facilities in 11 locations by 9 companies. In the US, EtO is primarily produced in Texas and Louisiana ("Ethylene Oxide Frequently Asked Questions," 2018).

Based on the 2014 National Emissions Inventory (NEI; https://www.epa.gov/air-emissions-inventories/2014-national-emissions-inventory-nei-data), Texas emits approximately 36% of the EtO in the US. As a result, tons of EtO emitted per square mile in Texas (1.8E-04 tons/square mile) is over 5 times higher in Texas compared to the rest of the US (3.5E-05 tons/square mile). Despite this and the extraordinarily high carcinogenic potency purported by USEPA (2016) for lymphoid and breast cancers, the incidences of leukemia and non-Hodgkin's lymphoma (both included in USEPA's dose-response assessment) as well as breast cancer are lower in Texas than in the general US population, with the same being true for all cancers combined (Table 3). Moreover, *Texas incidence rates are statistically significantly lower than the US for non-Hodgkin's lymphoma, breast cancer, and all cancer sites combined, with leukemia being of borderline statistical significance*. Again, leukemia, non-Hodgkin's lymphoma, and breast cancer are endpoints included in USEPA's carcinogenic dose-response assessment for EtO (USEPA 2016), along with multiple myeloma (state-specific versus US data were not available).

Table 3: Relevant Age-Adjusted US and Texas Cancer Incidence Rates per 100,000 (2012-1016)

Area	NEI Emissions (tons)	Emissions per Area (tons/mile^2)	Non-Hodgkin's Lymphoma	Leukemia	Breast Cancer (female)	All Cancer Sites
US	133.72	3.52E-05	19.2 (19.1, 19.3)	14.1 (14.1, 14.2)	125.2 (124.9, 125.4)	448.0 (447.7, 448.4)
Texas	48.45	1.80E-04	17.4 (17.2, 17.6)	13.9 (13.7, 14.1)	111.9 (111.2, 112.7)	407.7 (406.6, 408.9)

As to even more interesting examples regarding sources on a county level, USEPA's assessment of EtO as a potent carcinogen suggests that elevations in EtO-induced cancers should be expected in counties with relatively high EtO emissions per square mile and a sufficiently large population. With this in mind, the TCEQ notes that although highly-industrialized Jefferson County (population ≈260,000) has more EtO emissions on a square mile basis than any other county in Texas (1.1E-02 tons/square mile) with over 300 times more than the US at large (3.5E-

05 tons/square mile), the incidences of leukemia (13.4 (95% CI of 11.5, 15.5)), non-Hodgkin's lymphoma (17.5 (95% CI of 15.3, 19.9)), and breast cancer (102.4 (95% CI of 94.9, 110.3)) are lower in Jefferson County than in the general US population. *In fact, breast cancer incidence is statistically significantly lower in Jefferson County compared to both Texas and the US (see Table 3 for Texas and US rates), despite EtO emissions that are 60 times higher than Texas at large and 307 times higher than the US*. Based on USEPA's 2016 assessment, the opposite of this reality would be expected. Total cancer in Jefferson County (399.9 (95% CI of 389.3, 410.7)) is also statistically lower than in the US. Similarly, as by far the most populated Texas county (≈4.6 million) with relatively high reported NEI EtO emissions per square mile (i.e., 6.6E-03 tons/square mile is ≈188 times higher than the US at 3.5E-05 tons/square mile), the *incidences of leukemia* (13.0 (95% CI of 12.5, 13.5)), *non-Hodgkin's lymphoma* (16.9 (95% CI of 16.3, 17.5)), *breast cancer* (111.9.4 (95% CI of 109.9, 114.0)), *as well as all cancers combined* (400.1 (95% CI of 397.2, 403.1)) *are all statistically significantly lower in highly-industrialized Harris County than in the general US population*. Despite the associated uncertainties, such results may be viewed as both surprising and quite intriguing when considered in the context of USEPA (2016).

After the release of USEPA's 2014 National Air Toxics Assessment (NATA), the USEPA began to evaluate facilities that emit EtO. The 2014 NATA estimated that EtO significantly contributes to potential elevated cancer risks in some census tracts across the US; risk that is largely driven by the USEPA's recently-derived URF. Because of concerns related to cancer risk from EtO emissions raised by NATA, based on available information two EtO sterilizing facilities closed and two suspended operations in 2019. In order to prevent shortages of critical medical equipment, the US Food & Drug Administration (USFDA) has been working with medical device manufacturers to find alternative locations and methods for sterilization. According to the USFDA, EtO is the likely sterilant for medical devices made from certain polymers (plastic or resin), metals, or glass, or that have multiple layers of packaging or hard-to-reach places (e.g., catheters). Approximately fifty percent of all sterile medical devices in the US are sterilized with EtO ("Ethylene Oxide Sterilization,"2019).

Sources of EtO emissions into the air include, but are not limited to, industrial emissions or venting with other gases. Other sources of EtO air emissions include its use as a sterilizer of medical equipment and its release from commodity-fumigated materials. The general population may be exposed to EtO through breathing ambient air containing EtO, smoking tobacco products, and breathing secondhand cigarette smoke ("Ethylene Oxide. 75-21-8"). Certain occupational groups (e.g., workers in EtO manufacturing or workers that use EtO to produce solvents, antifreeze, textiles, detergents, and polyurethane foam, sterilization technicians, and agricultural workers involved in fumigation) may be exposed in the workplace (IARC 2012).

EtO is also produced endogenously in the body due to oxidation of ethylene, which is generated by intestinal bacteria, lipid peroxidation of unsaturated fats, methionine, and hemoglobin. Recent analyses indicate that endogenous levels of EtO are significant relative to (i.e., higher than) doses corresponding to recently-derived regulatory values (USEPA 2016) and typical environmental exposures (Kirman and Hays 2017).

Chapter 3 Carcinogenic Potential

3.1 Carcinogenic Weight of Evidence (WOE)

EtO has been evaluated for carcinogenic potential by the International Agency for Research on Cancer (IARC), the US Environmental Protection Agency (USEPA), and the World Health Organization (WHO). These agencies' carcinogenic classifications for EtO are provided in Table 4 below.

Table 4: Carcinogenic Weight of Evidence

Group	Classification
IARC (2012)	Group I: Carcinogenic to humans
USEPA (2016)	Carcinogenic to humans
WHO (2003)	Highly likely to be carcinogenic to humans

Generally, the TCEQ only performs carcinogenic dose-response assessments for chemicals considered by the TCEQ either to be "Carcinogenic to Humans" or "Likely to Be Carcinogenic to Humans" (TCEQ 2015). For the purposes of this DSD, the TCEQ has adopted the USEPA (2016) carcinogenic classification for EtO, which is discussed further in Section 3.3.2.

3.2 Relevant Data

3.2.1 Epidemiological Studies

Based on the systematic review conducted by the TCEQ (Appendix 1) as well as review of USEPA (2016) and other dose-response assessments (e.g., Valdez-Flores et al. 2010, Kirman et al. 2004), the assessment of excess cancer risk in the NIOSH and/or UCC cohorts provides the best basis for a carcinogenic assessment of EtO. These studies are summarized below.

3.2.1.1 NIOSH Cohort

Much of the following study summary is based on information from Section 4.1 of USEPA (2016).

The NIOSH retrospective cohort study of 17,530 workers in 13 sterilizing facilities (most recent follow-up by Steenland et al. 2004, 2003) provides adequate data for deriving quantitative cancer risk estimates (unit risk estimates or URFs) for EtO in humans. Briefly, the following are positive study attributes:

- Exposure estimates were derived for the individual workers using a comprehensive exposure assessment (although there are associated uncertainties);
- The cohort was large and diverse (e.g., 55% female); and
- There was little reported exposure to chemicals other than EtO.

EtO exposure estimates, including estimates for early exposures for which no measurements were available, were determined using a regression model that estimated exposures for each individual as a function of facility, exposure category, and time period. The regression model was based on extensive personal monitoring data from 18 facilities from 1976 to 1985 as well as information on factors influencing exposure, such as engineering controls (Hornung et al. 1994). Uncertainties are inevitably associated with historical exposure reconstruction. In this case, USEPA acknowledges that EtO measurement data were not available for most of the time that the cohort was exposed, errors in retrospective exposure assignments are inevitable, and that the exposure estimates are a primary source of uncertainty in the URF estimates (pp. 4-64 and 4-65 of USEPA 2016). Accordingly, to the TCEQ there appears to be appreciable uncertainty stemming from the lack of EtO exposure data prior to the time air monitoring data collection began when exposures for much of the cohort would have been relatively high and significantly contributed to cumulative exposure estimates (ppm-days, both unlagged and lagged), which appear likely to be biased low although a detailed discussion is beyond the scope of this DSD (e.g., Bogen et al. 2019, Li et al. 2019). The USEPA SAB agreed that these exposure estimates are likely of lower reliability (because there were no exposure measurement data that could be included in the exposure model prior to 1979) and actual EtO exposures were likely to have been higher than reflected in the estimates (p. I-41 of USEPA 2016). However, for the later monitoring data the regression model was able to account for 85% of the variation in average EtO exposure levels when evaluated against independent test data from the same set of data. The investigators estimated the cumulative exposure (ppm × days) for each individual worker by multiplying the estimated exposure (ppm) for each job (exposure category) held by the worker by the number of days spent in that job and summing over all the jobs held by the worker.

The TCEQ notes that this worker population was exposed to extraordinarily high concentrations of EtO. For example, Tables IV and V of Hornung et al. (1994) provide measured and estimated worker exposure means of 3.5-4.6 ppm, which are ≈1,000,000-2,000,000 times higher than central tendency environmental levels (i.e., background and environmental exposure means ≈0.0024-0.0034 ppb per USEPA 2016). Animal carcinogenicity studies were conducted at even

higher EtO exposure concentrations (10-100 ppm; see Section 3.2 of USEPA 2016). On any given day, estimated exposure for a job could have ranged from 50-77,000 ppb (pp. D-4 and D-37 of USEPA 2016), which is remarkably ≈15,000-32,000,000 times higher than central tendency environmental levels of EtO. Consequently, when USEPA (2016) discusses model fit in the "low-dose" region, the low-dose region for these workers provides no information about the shape of the dose-response at environmental levels. High-dose carcinogenicity data alone are incapable of informing truly low-dose risk, no matter how extensive the analyses or peer review (i.e., other relevant information must be duly considered).

In regard to study findings, Steenland et al. (2004) present follow-up results for the cohort mortality study previously discussed by Steenland et al. (1991) and Stayner et al. (1993). Findings in the most current follow-up include statistically increased (lympho)hematopoietic cancer mortality (i.e., non-Hodgkin's lymphoma with a 10-year exposure lag, haematopoietic cancer and lymphoid cell line tumors with a 15-year lag) in males but not females of the highest EtO exposure group (see Tables 4, 6, and 7 of the study), and statistically increased breast cancer mortality in females of the highest EtO exposure group with a 20-year lag but not without (see Tables 5 and 8 of the study). Steenland et al. (2003) present results of a breast cancer incidence study of a subcohort of 7,576 women from the NIOSH cohort that showed statistically increased odds ratios for the highest exposure group with a 15-year lag but not without (see Tables 4 and 5 of the study). No statistically significant increases in breast cancer were found for any exposure group using external referents and either 0- or 15-year exposure lags (see Table 3 of the study). These Steenland et al. studies were included in recent scientific literature reviews and meta-analyses of EtO studies for these cancer endpoints (Marsh et al. 2019, Vincent et al. 2019). See Appendix 7 for a discussion of the overall weight of evidence (across studies) concerning EtO-induced breast cancer in humans.

3.2.1.2 UCC Cohort

Swaen et al. (2009) redefined and updated the UCC cohort of male workers employed in industrial facilities where EtO was produced or used. Previous studies of the UCC cohort were published by Greenberg et al. (1990) and Teta et al. (1993). All 2,063 men were employed between 1940 and the end of 1988 and were observed for mortality through 2003. Workers from EtO departments at the Kanawha Valley, WV sites hired after 1988 were determined to have no appreciable EtO exposure and were, therefore, not added to the cohort. Cause-specific standardized mortality ratios (SMRs) were calculated. Internal analyses were made by applying Cox proportional hazards models to the data.

The exposure assessment for this update relies on the qualitative categorization of EtO producing and using departments by exposure level developed by Greenberg et al. (1990), and on quantitative estimates of average intensity by these department categories and by time period (1925-1988) developed by Teta et al. (1993). Time period cut points were chosen as

follows: 1925, the start-up of EtO production in the Kanawha Valley; 1940, start of cohort observation and first period with published estimates of exposure; 1957, chlorohydrin process for EtO production completely shut-down; and 1974, the period when airborne exposures declined substantially due to process and exposure controls. The combination of the average exposure for the four different time periods and the three classifications of departments into low, medium, and high exposure departments created the exposure matrix. Cumulative EtO exposure (ppm-years) for each study subject was then estimated by multiplying the estimated time-period and department-specific exposure concentrations by duration in months for each individual's assignments to EtO departments and summing the products over all assignments up through December 1988 (Swaen et al. 2009). The average cumulative EtO exposure was 67.16 ppm-years (≈16,118 ppm-days, as 67.16 ppm-years × 240 days/year), about twice that of the NIOSH cohort. As of Swaen et al. (2009), the average follow-up period for the UCC cohort was 10 years longer (36.5 versus 25.8 years) and the percent deceased was 3-fold greater than the NIOSH cohort (51% versus 16%). However, the number of expected cancer deaths for the UCC cohort (a measure of study power) was between 2-3 times less because of the significantly smaller cohort size in both number and person-years (e.g., 75,306 versus 450,906 person-years). Nevertheless, this is an important cohort that contributes to the human EtO carcinogenicity database. For example, relatively high cumulative exposure estimates and a long follow-up period are often viewed as advantageous for the identification and complete ascertainment of chemically-induced cancers. In this case, however, USEPA (2016) indicates that the long follow-up may be viewed as a limitation as well since "the follow-up is likely observing workers at the high tail end of the distribution of latency times for EtO-associated lymphohematopoietic cancers", and characterizes the exposure assessment as relatively crude (see Section A.2.20 of USEPA 2016).

As mentioned above, uncertainties are inevitably associated with historical exposure reconstruction. In addition to finding fault with the cohort for being smaller and limited to males, USEPA (2016) characterizes the EtO exposure assessment for the UCC cohort as more uncertain than that for the NIOSH cohort (e.g., greater likelihood for exposure misclassification, use of surrogate exposure data; see Section 4.1 of USEPA 2016). USEPA further indicates that there are significant uncertainties in the exposure estimates for the early years when the highest exposures occurred (Section A.2.20 of USEPA 2016), something both cohorts apparently have in common. That is, the NIOSH cohort appears to have unemphasized yet significant uncertainties of its own; most notably, the lack of exposure data prior to the mid-70's when exposures were likely to have been significant and would have increased cumulative exposure estimates for much of the cohort (e.g., Bogen et al. 2019). Ultimately, the TCEQ finds that these cohorts are the best candidates for regulatory dose-response assessment for EtO, and cohort-specific results will be appropriately weighed based on relevant statistical criteria.

Regarding study findings, Swaen et al. (2009) report that no indications were found for excess cancer risks from EtO exposures, including the lymphohematopoietic malignancies (e.g., 11 leukemia deaths occurred and 11.8 were expected, 12 non-Hodgkin's lymphoma deaths occurred and 11.5 were expected). Cox proportional hazards modeling for all cause, leukemia, and lymphoid malignancies mortality revealed no trends or associations with cumulative EtO exposure. In recognition of exposure estimate uncertainty, it is also important to note that no statistically significantly elevated SMRs were found in the analysis by hire date, and there were no statistically significant increases in the longest duration category and no suggested trends by duration (all surrogates of exposure). Study authors concluded that the cohort showed no long-term carcinogenic effects associated with EtO exposure.

Similarly, an as of yet unpublished update of the UCC cohort through 2013 (submitted as Bender et al.) concludes that examination of mortality from all causes of death, all cancers, leukemia, non-Hodgkin's lymphoma, and lymphoid malignancies revealed no evidence for an exposure-related response; EtO exposure in this cohort was not associated with an observable increase in lymphohematopoietic cancer mortality (personal communication with Ciriaco Valdez-Flores, an author of a risk assessment paper based in part on the Bender et al. update). The average cumulative dose of EtO (67 ppm-years) is reported to be around two times that for the NIOSH cohort, with a ≈63% longer follow-up period (≈41 years) and a similar number of lymphoid cancer deaths in males (27 in NIOSH versus 25 in UCC) despite the number of person-years for males in the NIOSH cohort (189,868 person-years) being significantly greater than that in the UCC cohort (83,524 person-years).

3.2.2 Animal Studies

Human (i.e., epidemiological) data are available for a carcinogenic assessment of EtO and are preferred over animal data for toxicity factor (i.e., URF) development (TCEQ 2015). Therefore, animal carcinogenicity data are not discussed herein in regard to dose-response assessment (see Section 4.2 of USEPA 2016 for relevant information). However, laboratory animal carcinogenicity data for EtO must be discussed in regard to their ability to support human data, which USEPA (2016) acknowledges are insufficient alone to establish that EtO is *carcinogenic to humans*.

Based on their respective URFs, USEPA estimates that EtO is ≈1,000-fold more carcinogenic than benzene. Given such a high carcinogenic potency purported by USEPA (2016) combined with the large groups of workers (including women) exposed to very high concentrations of EtO on a daily basis (up to tens-of-millions of times higher than the ambient levels cited in USEPA 2016), the fact that human data alone are admittedly insufficient to classify EtO as *carcinogenic to humans* is quite extraordinary. As a result, USEPA must rely on support by laboratory animal studies in classifying EtO as *carcinogenic to humans*. However, upon closer scrutiny of the underlying science, the sites of EtO-induced cancers in animal models are of questionable

human relevance for being predictive of, or confirming evidence for, the site(s) of human cancers. As a cancer endpoint of interest, breast cancer is used as one of the examples below.

3.2.2.1 Data Relevance for Human Cancer Site Concordance

While laboratory animal data are often used to support various aspects of regulatory assessments, interspecies differences in carcinogenic responses are common (e.g., tumors types, sensitivity), even between rodents (e.g., EtO induced mammary tumors in mice but not rats). For example, IARC (2019) analyzed tumor site concordance using a dataset of the 111 distinct Group 1 (*carcinogenic to humans*) agents identified up to and including Volume 109. Sixty agents had both a human tumor site and an animal tumor site identified and were used to evaluate concordance across 39 tumor sites in animals and humans (see Figures 21.1 and 21.2 of IARC 2019). Reported results show *that breast cancer is more frequently/commonly induced in laboratory animal species* than in humans. More telling is that while there is 47% overlap between agents that cause lymphoid and haematopoietic cancers in humans and animals, *there is only 20% overlap between agents that have been shown to cause breast cancer in humans and animals* (Table 21.7 of IARC 2019). The IARC (2019) unanimous consensus statement [*emphasis added*] is that "At present, *the state of the science does not support tumour site concordance as a general principle*."

Accordingly, current best available science indicates that animal data cannot generally be used to support specific sites of chemically-attributable carcinogenesis in humans; even more so when laboratory animal results are inconsistent (e.g., mammary tumors in mice but not rats) and the human database is relatively robust. For example, EtO-induced *murine mammary tumors* are not even predictive for rats. Additionally, while *lung cancer* was statistically increased in both male and female mice at incidences of 53% and 45%, respectively (Table 3-3 in USEPA 2016), lung cancer is a not a candidate endpoint in humans as *data for this very strong carcinogenic response in mice are simply not predictive for humans* (i.e., no interspecies site concordance; SMR of 1.05 (0.95, 1.17) in Table 1 of Steenland et al. 2004). Similarly, EtO induced statistically significant increases in *brain tumors* in rats of both sexes (Table 3-5 in USEPA 2016), *but yet again, these results are not predictive for humans*. In fact, *brain cancer for the NIOSH cohort is statistically significantly decreased* (i.e., SMR of 0.59 (0.36, 0.91) in Table 1 of Steenland et al. 2004), just the opposite of what the rat data would suggest. *Clearly, laboratory animal data for EtO-induced cancers cannot be relied upon to identify cancer sites or otherwise predict EtO carcinogenic response in humans*. This applies to cancer sites generally and EtO-induced breast cancer specifically since: (1) The state of the science does not support tumor site concordance as a general principle (IARC 2019); (2) Specific to breast cancer, there is relatively little overlap between agents that have been shown to cause breast cancer in humans and animals (i.e., there are significant interspecies differences), with *discordance* generally being the case (IARC 2019); and (3) Specific to EtO, animal data are simply not reliable

predictors of the purported sites of EtO-induced carcinogenesis in humans (e.g., lung and brain cancer in laboratory animals).

3.3 Mode of Action (MOA) and Carcinogenic Classification

The TCEQ has adopted the overall USEPA (2016) EtO MOA analysis and carcinogenic classification determinations for the purposes of this DSD. For example, USEPA's carcinogenic classification allows for a carcinogenic dose-response assessment under TCEQ guidelines (TCEQ 2015). This does not necessarily mean, however, that the TCEQ necessarily fully concurs with every USEPA statement or characterization. As such, summary information was essentially derived directly from Sections 3.4.3 and 3.5.1 of USEPA (2016) and is presented below, with references to USEPA (2016) document sections removed [*emphasis added*]. The references for the studies supporting the information below can be found in the aforementioned sections of USEPA (2016).

3.3.1 MOA

In this section, the evidence for a mutagenic MOA for EtO carcinogenicity is analyzed under the MOA framework in the USEPA's 2005 Guidelines for Carcinogen Risk Assessment (USEPA 2005a).

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

Is the hypothesized MOA sufficiently supported in the test animals?

Consistent with the USEPA's 2005 Guidelines for Carcinogen Risk Assessment (USEPA 2005a), this MOA analysis for a mutagenic MOA is organized around the Hill "criteria" (or considerations) developed for the analysis of epidemiological studies (Hill 1965). These considerations are denoted in *underlined italics* in the discussion below.

Numerous studies have demonstrated that EtO forms protein and DNA adducts, in mice and rats, and there is incontrovertible evidence that EtO is mutagenic and genotoxic. The evidence for causal associations between the key events and tumor formation has *strength* and *consistency*. Increases in the frequency of gene mutations in reporter genes have been observed in the lung, T-lymphocytes, bone marrow, and testes of transgenic mice and in T-lymphocytes of rats exposed to EtO via inhalation at concentrations similar to those inducing

tumors in the rodent carcinogenesis bioassays. In addition, in the lung, uterine, mammary gland and Harderian gland tumors from EtO-exposed mice in those bioassays, dramatic shifts toward guanine and adenine mutations have been observed in the mutational spectra of the proto-oncogenes *Hras* and *Kras*, as well as the tumor suppressor *Trp53*, consistent with the propensity of EtO to form DNA adducts on purine bases.

Inhalation studies in laboratory animals have also demonstrated that EtO exposure levels in the range of those used in the rodent bioassays induce sister chromatid exchanges (SCEs) in several species and consistently induce chromosomal aberrations in mice. In rats, although SCEs are consistently observed in the available studies, the results for micronuclei formation and chromosomal aberrations following subchronic (up to 4-week) inhalation exposures to EtO at the same exposure levels as those used in the rodent bioassays have been nonpositive; however, IARC (2008) has noted analytical limitations with some of these analyses. In addition, Donner et al. (2010) demonstrated a clear duration effect in mice, with chromosomal aberrations being induced at those same EtO exposure levels only following longer exposure durations (≥12 weeks).

Specificity is not expected for a multisite mutagen and carcinogen such as EtO (USEPA 2005a). A *temporal relationship* is clearly evident, with DNA adducts, point mutations, and chromosomal effects observed in acute and subchronic assays.

Dose-response relationships have been observed between EtO exposure *in vivo* and DNA adducts, SCEs, and *Hprt* and *Trp53* mutations. A mutagenic MOA for EtO carcinogenicity also clearly comports with notions of *biological plausibility* and *coherence* because *EtO is a direct-acting alkylating agent. Such agents are generally capable of forming DNA adducts, which in turn have the potential to cause genetic damage, including mutations; and mutagenicity*, in its turn, is a well-established cause of carcinogenicity. This chain of key events is consistent with current understanding of the biology of cancer.

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

Is the hypothesized MOA relevant to humans?

The evidence discussed above demonstrates that EtO is a systemic mutagen in test animals; thus, there is the presumption that it would also be a mutagen in humans. Moreover, human evidence directly supports a mutagenic MOA for EtO carcinogenicity. Several studies of humans have reported exposure-response relationships between hemoglobin adduct levels and EtO exposure levels (e.g., van Sittert et al. 1993, Schulte et al. 1992), demonstrating the ability of

EtO to bind covalently in systemic human cells, as it does in rodent cells. DNA adducts in EtO-exposed humans have not been well studied, and the evidence of increased DNA adducts is limited. EtO has yielded positive results in *in vitro* mutagenicity studies of human cells. Although the studies of point mutations in EtO-exposed humans are few and insensitive and the evidence for mutations is limited, there is clear evidence from a number of human studies that EtO causes chromosomal aberrations, SCEs, and micronucleus formation in peripheral blood lymphocytes, with some evidence of positive relationships with exposure concentration and duration.

USEPA (2016) concludes that *the weight of evidence supports a mutagenic MOA for EtO carcinogenicity*. Although oxidative stress or other processes might contribute to the development of EtO-induced cancers, the TCEQ agrees that the available evidence best supports a mutagenic MOA as the primary process mediating EtO-induced carcinogenicity (USEPA 2016).

3.3.2 Carcinogenic Classification

Regarding carcinogenic classification under USEPA (2005a), while USEPA (2016) states that there is substantial evidence that EtO exposure is causally associated with lymphohematopoietic cancers and female breast cancer in human studies, the agency acknowledges that *the evidence is not strong enough to be conclusive*. Of the seven relevant Hill "criteria" (or considerations) for causality (Hill 1965), *temporality*, *coherence*, *biological plausibility*, and *analogy* are readily satisfied, and the other three criteria (*consistency*, *biological gradient*, and *strength of association*) are satisfied to varying degrees. For example, there is "*some* evidence of dose-response relationships", but "there is *little strength* in the associations, as reflected by the modest magnitude of most of the RR [relative risk] estimates." See Section 3.5.1 of USEPA (2016) for additional discussion on these criteria. USEPA (2016) ultimately concludes that *the overall epidemiological evidence for causal associations between EtO exposure and lymphohematopoietic cancer as well as female breast cancer is strong but less than conclusive, with epidemiology study evidence for other cancer types (e.g., stomach cancer and pancreatic cancer) also being inadequate*. The TCEQ particularly disagrees with this characterization of the strength of evidence for breast cancer (e.g., see Appendix 7). The results of recent meta-analyses are consistent with TCEQ's conclusions as to breast cancer, and further disagree with USEPA's characterization of the evidence for lymphohematopoietic cancer (Marsh et al. 2019 and Vincent et al. 2019 are discussed briefly below).

USEPA (2016) further concludes that the laboratory animal evidence for EtO carcinogenicity is "sufficient" based on findings of tumors at multiple sites, by both oral and inhalation routes of exposure, and in both sexes of both rats and mice. Tumor types resulting from inhalation exposure included mononuclear cell leukemia in male and female rats and malignant lymphoma and mammary carcinoma in female mice. Note, however, that IARC (2019)

concluded that the state of the science does not support tumor site concordance between humans and animals as a general principle, and concordance was particularly poor for mammary tumors (i.e., IARC's analysis demonstrates that discordance between humans and laboratory animals is the general rule for mammary tumors based on available data). Lastly, USEPA concludes that the evidence of EtO genotoxicity and mutagenicity is unequivocal (see Section 3.5.1 of USEPA 2016 for details), ultimately classifying EtO as "*carcinogenic to humans*" under USEPA (2005a).

The TCEQ agrees that since the epidemiological evidence is less than convincing, additional lines of evidence are required for the EtO carcinogenic classification. USEPA (2016) cites the following lines of evidence to support the "*carcinogenic to humans*" classification: (1) there is strong, although less than conclusive on its own, evidence of cancer in humans associated with EtO exposure via inhalation, specifically, evidence of lymphohematopoietic cancers and female breast cancer in EtO-exposed workers; (2) there is extensive evidence of EtO-induced carcinogenicity in laboratory animals, including lymphohematopoietic cancers in rats and mice and mammary carcinomas in mice following inhalation exposure (*see Section 3.2.2.1 of this DSD for discussion regarding interspecies site concordance issues*); (3) EtO is a direct-acting alkylating agent whose mutagenic and genotoxic capabilities have been well established in a variety of experimental systems, and a mutagenic mode of carcinogenic action has been identified in animals involving the key precursor events of DNA adduct formation and subsequent DNA damage, including point mutations and chromosomal effects; and (4) there is strong evidence that the key precursor events are anticipated to occur in humans and progress to tumors, including evidence of chromosome damage, such as chromosomal aberrations, SCEs, and micronuclei in EtO-exposed workers. In supporting a "*carcinogenic to humans (Group 1)*" designation, IARC (2012) draws conclusions similar to those of USEPA (2016), citing *limited evidence* in humans, *sufficient evidence* in experimental animals, and *strong evidence* supportive of a genotoxic MOA for carcinogenicity.

As mentioned in Section 3.3, the TCEQ has simply adopted the USEPA (2016) EtO carcinogenic classification (*carcinogenic to humans*) for purposes of this DSD, which allows for a carcinogenic dose-response assessment under TCEQ guidelines (TCEQ 2015). However, recent publications call into question USEPA's characterization of the strength of the evidence as well as their classification. Specifically, the meta-analyses and other information in Marsh et al. (2019) and Vincent et al. (2019) raise serious questions about the accuracy of USEPA's characterization of the overall epidemiological evidence for EtO-induced lymphohematopoietic cancer and breast cancer as strong. Vincent et al. (2019) further indicate that, "Similarly, toxicological studies and studies of early effect biomarkers in animals and humans provided no strong indication that EtO causes LHM or mammary cancers", and conclude that… "the IARC and USEPA classification of EtO as a known human carcinogen overstates the underlying evidence." As mentioned above, the TCEQ particularly disagrees with USEPA's characterization of the strength of evidence for

breast cancer (e.g., see Appendix 7), for which Vincent et al. (2019) conclude that the evidence from several studies of workers exposed to relatively high concentrations of EtO over relatively long durations in a range of workplace settings fails to demonstrate clear or consistent associations with occupational exposure to EtO. Ultimately, the study authors conclude that: (1) the epidemiological evidence certainly does not comport with USEPA's conclusion that the evidence was "strong" (i.e., the broad body of epidemiological studies demonstrates no increased cancer risks); (2) the available *in vivo* 2-year cancer bioassays provide only limited evidence that EtO causes mammary tumors (or LHM cancers); and (3) integrating the genotoxic MOA with the epidemiological literature and animal toxicology evidence aligns most closely with a *suggestive evidence of carcinogenic potential* classification for EtO. Vincent et al. nevertheless acknowledge that a quantitative dose-response assessment may be warranted based on the likely MOA.

In conclusion, USEPA (2016) acknowledges that human data are insufficient to demonstrate that EtO causes any particular cancer in humans (e.g., lymphoid or breast cancer); and this in workers exposed to levels up to millions of times higher than environmental levels to which the general public may be exposed. Additionally, IARC's recent conclusion that the state of the science does not support tumor site concordance between humans and animals as a general principle casts serious doubt on USEPA using animal results to support any particular cancer endpoints, especially mammary tumors (as the agency did) for which interspecies concordance is particularly poor (IARC 2019). Nevertheless, this assessment evaluates both lymphoid cancer and breast cancer as candidate endpoints. While the weight of epidemiological evidence for EtO-induced lympho-hematopoietic cancers may be viewed as somewhat greater than that for breast cancer (Section 3.4.1.6.1), the human evidence is still acknowledged by both USEPA (2016) and the TCEQ as insufficient. Considering the admittedly inconclusive human evidence for EtO-induced cancer in workers exposed long-term to extraordinarily high EtO concentrations, the derivation of carcinogenicity-based toxicity factors may be conservative.

3.4 Carcinogenic Dose-Response Assessment

Per TCEQ guidelines (TCEQ 2015), when toxicity factors or guideline air levels are identified in the scientific literature or databases, they are reviewed to determine whether the approaches used to develop the toxicity factors or guideline levels are similar to the procedures used by the TCEQ. If so, after careful consideration the TCEQ may elect to adopt the published toxicity factor or guideline level. In the present case, the scientific literature search identified USEPA (2016) and Valdez-Flores et al. (2010) as representing two relatively recent carcinogenic dose-response assessments for EtO for consideration under TCEQ guidelines (TCEQ 2015). In Sections 3.4.1.1, 3.4.1.2, and 3.4.1.3 below, the TCEQ reviews the available data for MOA, endogenous EtO levels, key epidemiological data, model predictions and fit criteria to determine whether to adopt the USEPA (2016) EtO inhalation URF. After making the adoption determination, the

TCEQ continues with an original assessment to derive an EtO inhalation URF based on TCEQ guidelines and best principles.

3.4.1 Low-Dose Extrapolation Approach

Use of MOA information to inform the dose-response assessment is a main focus of the TCEQ (2015) and USEPA (2005a,b) guidelines. Consequently, examining the MOA (as well as dose-response) for cancer endpoints with statistically significant increases (e.g., endpoint-specific SMRs) is an important initial step in cancer dose-response assessment. Generally, the MOA and other information may support one of the following low-dose extrapolation approaches: (1) nonthreshold (typically a linear extrapolation to zero); (2) threshold (typically identifying a point of departure (POD) and applying uncertainty factors); or (3) both (TCEQ 2015). Thus, to the extent that an MOA for a chemical is understood, it informs the low-dose extrapolation procedure for that chemical. Examples of different shapes of dose-response curves are shown in Figure 1.

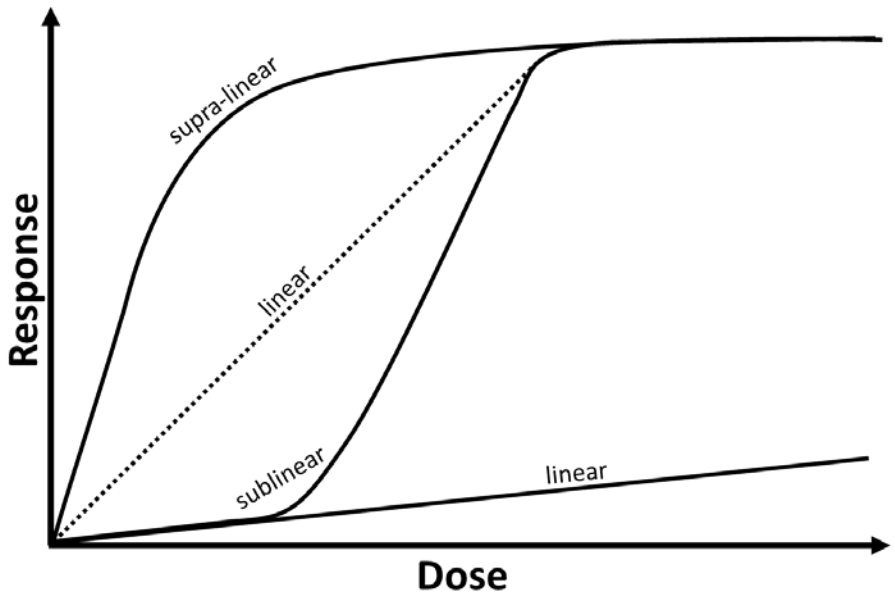


Figure 1: Dose-Response Curve Examples

One of the potential non-threshold dose-response curves shown in Figure 1 is supra-linear, with a steep low-dose slope. Section 7.7.7.3 of TCEQ (2015) addresses the potential use of a supra-linear dose-response model specifically, and indicates [*emphasis added*]:

> As indicated by Crump and Allen (1985), linear exposure-response models are "considered conservative in the sense that other biologically plausible dose-

> response models would generally imply lower risks." *Some researchers have published dose-response models that are inherently supra-linear at low exposures*. The increase of the hazard rate or relative risk of a supra-linear exposure-response model is faster at lower exposures than at higher exposures. *These types of models are generally not biologically plausible and tend to grossly exaggerate the estimation of risks at low exposures*.

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

In this evaluation, the TCEQ concludes that the low-dose extrapolation of EtO-induced carcinogenic effects should be based on a model that is no more than linear overall, and arguably sublinear at endogenous levels and below. This conclusion is based on data relevant to the MOA, data on normal background endogenous EtO levels, key epidemiological study data (e.g., overall results for the UCC and NIOSH cohorts, doses associated with statistically increased cancer), and an evaluation of the ability of a model (or lack thereof) to reasonably predict the underlying dose-response data. *In contrast, USEPA (2016) used an overall supra-linear two-piece spline model for EtO carcinogenesis. This model assumed a non-threshold steep linear relationship between EtO and cancer formation at lower concentrations, with a high-concentration linear relationship that had a much shallower slope (i.e., an overall supra-linear relationship; similar to that depicted in Figure 1). In the present case, the TCEQ finds insufficient data to justify the supra-linear modeling approach (i.e., use of the steep lower-dose slope starting at zero dose from the linear two-piece spline model) ultimately adopted by USEPA (2016).* Even ignoring more critical and discerning considerations in this particular case (e.g., the lack of mechanistic data to justify use of an overall supra-linear model, model predictions of the underlying key cohort data), the appropriate consideration of model fit criteria (e.g., for lymphoid cancer mortality in the NIOSH cohort) still does not support use of a supra-linear model over more conventional models (e.g., the likelihood of the linear two-piece spline model for lymphoid cancer mortality is not different from the likelihood of the null model at the 5% significance level, visual examination of model fits to the actual underlying data, etc.). These considerations are discussed in more detail below.

3.4.1.1 Consideration of MOA

MOA information is discussed in Section 3.3.1, which supports a likely mutagenic MOA for EtO carcinogenicity. *MOA information can suggest the likely shape of the dose-response curve at lower doses (USEPA 2005a).* That is, toxicological principles can inform expectations about low-dose risk when truly low-dose data are unavailable. In this case, in the key epidemiological

cohort (NIOSH) used by USEPA (2016), estimated mean worker exposures to EtO were ≈1,000,000-2,000,000 times higher than central tendency ambient environmental EtO levels (see Section 3.2.1.1). Consideration of a direct acting DNA-reactive chemical in conjunction with normal detoxification processes and baseline levels of DNA repair enzymes that have evolved to efficiently detoxify and/or repair significant levels of endogenous EtO and associated adducts (in the endogenous range) suggests a no more than linear low-dose response component near the endogenous range with a transition to a higher dose-response slope at some point above the endogenous range where the body can no longer effectively detoxify EtO and/or repair the resulting damage. Thus, across a range of doses from truly low (e.g., ambient air, endogenous) to high (e.g., high occupational exposures), the expected dose-response could be characterized as appearing sublinear in the low-dose range and/or sublinear overall across doses (see Figure 1). In contrast to direct acting chemicals such as EtO, supra-linear responses are associated with an MOA that involves the saturation of metabolic activation where fewer electrophiles are formed per unit dose at higher exposures, which is not the case for EtO (Swenberg et al. 2008).

Kirman and Hays (2017) expressed this conclusion similarly. That is, based on relevant considerations, an overall sublinear (not supra-linear) dose-response would be expected over the range of possible exposures to EtO, from those that result in total body burdens (endogenous + exogenous) within the normal endogenous level range to those that result in a total body burden significantly greater than the normal range where the normally effective detoxification/repair processes are overwhelmed. This conclusion is reasonably consistent with that of the USEPA [*emphasis added*], "EPA considers it *highly plausible that the dose-response relationship over the endogenous range is sublinear* (e.g., that the baseline levels of DNA repair enzymes and other protective systems evolved to deal with endogenous DNA damage would work more effectively for lower levels of endogenous adducts), that is, that the slope of the dose-response relationship for risk per adduct would increase as the level of endogenous adducts increases." As equal internal doses give rise to equal risk as a matter of toxicological principal, the expectation of sublinearity also applies when total internal exposure (endogenous plus a relatively minor contribution from exogenous) falls within the range of normal endogenous background. Figure 4 and Figure 7 in Section 3.4.1.2.2.2 show that *EtO exposures corresponding to 1E-06 to 1E-04 excess risk based on USEPA (2016) are well below those corresponding to normal endogenous background levels*, inevitably leading to the expectation of sublinearity (or no excess risk) at such low doses based on the discussions above. In contrast to an overall linear or sublinear model, *using an overall supra-linear dose-response model (i.e., the steep low-dose component) to extrapolate risk down to an exposure lower than the point where a transition to a sublinear dose-response would be expected is not scientifically defensible and would be expected to grossly exaggerate truly low-dose risk (e.g., at endogenous levels and below).* That is, a steep slope from one portion of an overall supra-linear dose-response model should not be applied to a portion of the dose-response that admittedly is highly likely to have a shallow/sublinear slope.

Lastly, in addition to USEPA citing direct mutagenic activity as mechanistic justification for default linear extrapolation from high-to-low doses (pp. 4-22 and 4-37 of USEPA 2016) while still considering it "highly plausible that the dose-response relationship over the endogenous range is sublinear," *it is also critical to note that USEPA acknowledges the lack of mechanistic data to support the biological plausibility of an overall supra-linear dose-response, stating "the EPA is not aware of a mechanistic explanation" in response to questions from the USEPA SAB* (p. I-29 of USEPA 2016). Consistent with this acknowledgment, Vincent et al. (2019) consider the MOA and dose-response analysis of the early effect data in humans/animals (as well as modeling results of relevant cancer endpoints in rodents; most notably, leukemia incidence in female F344 rats) to conclude that there is no evidence that a dose-response other than linear is justified (e.g., "the USEPA's derivation of the IUR for EtO using a 2-piece, supralinear dose–response model… appears not to be adequately justified based on the published literature and deviates from USEPA standard risk assessment guidance"). Since lymphoid cancer drove the USEPA carcinogenic assessment, perhaps the most relevant mutagenicity data discussed by USEPA (2016) was that in the bone marrow of mice exposed to 25-200 ppm EtO by inhalation *in vivo* (Recio et al. 2004), which *USEPA indicates is consistent with a linear dose-response* (see C-17 of USEPA 2016), at least at doses well above endogenous and has not plateaued even at 200 ppm.

In summary:

- An overall sublinear dose-response is expected for endogenous, direct-acting chemicals like EtO where truly low dose (e.g., endogenous) to high dose response data are available (i.e., an overall more-than-/supra-linear dose-response is not expected).
- USEPA acknowledges that it is *highly plausible* that the EtO dose-response relationship over the endogenous range is *sublinear*, and since the exposures corresponding to 1E-06 to 1E-04 excess risk based on USEPA (2016) are *well below* those corresponding to normal endogenous background levels (see Figure 4 and Figure 7 in Section 3.4.1.2.2.2), a sublinear dose-response would be expected at such low doses (if any biologically meaningful response is to be expected).
- Consequently, *it is not scientifically defensible and likely grossly exaggerates EtO low-dose risk to use a supra-linear dose-response model (i.e., the steep low-dose slope) to extrapolate risk below the point where a transition to a sublinear dose-response is expected or "highly plausible" (e.g., at endogenous levels and below)*.
- As body burdens progressively increase to significantly higher levels than the normal endogenous range, excess risk is expected to increase as the normally relatively effective detoxification/repair processes are progressively overwhelmed at higher and higher doses, *making higher-than-endogenous risk increasingly discernable from background risk consistent with the assessment of "excess" (i.e., above background) risk*.

The USEPA should not have used an overall supra-linear model (i.e., the linear two-piece spline model) to derive a URF for EtO without a robust mechanistic justification for expecting supra-linearity (i.e., the steep lower-dose slope component) at truly low doses or used it to make a large low-dose extrapolation through and below an area (i.e., the endogenous range) where the agency actually considers sublinearity as "highly plausible." The NIOSH data are, in fact, not inconsistent with such expectations at low doses of EtO as there are no truly low-dose data available from the cohort (Section 3.2.1.1).

3.4.1.2 Consideration of Endogenous Levels, Key Epidemiological Data, and Model Predictions

3.4.1.2.1 Endogenous Levels

Considering its genotoxicity and relatively high occupational exposure levels, Coggon et al. (2004) consider the relatively low cancer risk associated with occupational exposure to EtO to be somewhat surprising and suggest the explanation may lie in the capacity of human cells to repair the DNA damage caused by EtO, which also occurs naturally through the action of endogenously formed EtO. The analysis of Kirman and Hays (2017) documents endogenous EtO levels normally found within the body expressed in terms of exogenous EtO exposures. Such information can provide support for a given risk assessment approach for chemicals such as EtO that have both endogenous and exogenous exposure pathways. Hemoglobin N-(2-hydroxyethyl)-valine (HEV) adducts are caused by the reaction of EtO with hemoglobin in erythrocytes and provide a biomarker/molecular dosimeter of internal EtO dose that can be correlated with exogenous (i.e., ambient air) EtO exposure. USEPA (2005a) indicates that it may be informative to use such biomarkers of internal exposure for dose-response assessment or to provide insight into the potential shape of the dose-response curve at doses below those at which tumors are induced experimentally. As EtO is widely distributed in the body, the levels of HEV in erythrocytes are expected to be proportional to levels of HEV in other tissues (including target tissues), which are further expected to be proportional to tissue exposures to free EtO (Kirman and Hays 2017). Kirman and Hays (2017) conducted a meta-analysis from the published literature characterizing the distribution of HEV adducts in EtO-unexposed (i.e., the background endogenous distribution) and exposed populations (smokers, workers). The relationship between exposure and HEV adducts is linear with R^2=0.998 (see Figure 3 of the study). In the meta-analysis for unexposed populations (n=661), the weighted mean of background endogenous HEV (random effects model) was 21.1 pmol/g Hb with a standard deviation (SD) of 14.6 pmol/g Hb. The fixed effects model produced very similar results (see Table 3 of the study).

The TCEQ notes that the reported mean human background endogenous HEV level of 21.1 pmol/g Hb appears reasonable given background HEV levels in control rats (≈42-50 pmol/g Hb) and mice (≈58-100 pmol/g Hb) (Walker et al. 1993, 2000). Furthermore, exposure to typical

environmental levels (i.e., background and environmental exposure means ≈0.0024-0.0034 ppb per USEPA 2016) would not be expected to substantially affect Kirman and Hays (2017) estimates of endogenous levels since they are well below the continuous air concentration corresponding to even the first percentile of the distribution (i.e., the 1st percentile of the distribution corresponds to a continuous air concentration of ≈0.37 ppb). Thus, within this document these data are simply referred to as endogenous.

The air concentrations corresponding to various endogenous level summary statistics (e.g., mean, 5th and 95th percentiles) from Kirman and Hays (2017) are able to provide valuable context for exogenous exposure concentrations. That is, considering the normal range of endogenous levels provides some important context for exogenous exposures and the likelihood that they may be biologically meaningful. For example, the TCEQ notes that ≈1.9 ppb is the continuous EtO air exposure concentration that corresponds to the endogenous EtO mean, and ≈4.5 ppb is the air concentration that corresponds to the 95th percentile of the endogenous distribution (Table 4 of Kirman and Hays 2017). Additionally, ≈1.3 ppb is the continuous exposure level that corresponds to an endogenous EtO increase of 1 SD (corresponding to an HEV increase of 14.6 pmol/g Hb). This does not provide a basis for expecting a biologically meaningful increase in internal exposure for this endogenous chemical with continuous exogenous exposure to less than 1 ppb EtO when considering the normal endogenous range (e.g., a continuous exogenous exposure of 2.4 ppb would be required to move those at the 95th percentile to the 99th percentile). *More specifically, this suggests that inhalation exposure to sub-ppb EtO air concentrations, particularly concentrations in the range of parts per trillion (e.g., 0.1-10 ppt), is of little biological importance compared to normal endogenous background levels.* For example, continuous exposure to 1 ppt EtO (i.e., the 1 in 100,000 excess risk air concentration using USEPA's URF) would result in ≈0.0109 pmol/g Hb added HEV, a mere 0.075% of the SD for normal background endogenous levels and over 360 times less than even the 1st percentile of normal background endogenous levels. This magnitude of change in HEV may be reasonably characterized as biologically insignificant. Considering EtO as a mutagenic carcinogen, this is generally consistent with the conclusions of Swenberg et al. (2008) that:

1) At low exposures, the likelihood that a mutation will arise from exogenous adducts becomes *de minimus* as compared to the large molecular dose normally formed endogenously (e.g., *Hprt* mutations were not statistically increased compared to background in mice exposed even up to 10 ppm EtO; high exposure to ≥50 ppm EtO was required to produce statistically significant increases over background; see Figure 9 of the study); and
2) The biologic effects of *de minimus* exposures below endogenous amounts are lost in the noise of the background (e.g., carcinogenesis is driven by endogenous DNA damage

when the dose-response for mutations due to external EtO exposure comes into the normal background frequency due to endogenous production).

Put another way by Kirman and Hays (2017), pragmatically speaking, the considerable variation in endogenous EtO exposure creates a signal-to-noise issue when exogenous exposures fall well below those consistent with endogenous exposures, and in such cases small exogenous exposures may not contribute to total exposure or to potential effects in a biologically meaningful way. Note that dose-response modeling for the actual carcinogenic endpoint(s) of interest is conducted later in this DSD, and that information on endogenous levels is only meant to provide some additional context for risk-based results and does not play a key role in model selection (e.g., unlike MOA, followed by model predictions of the underlying key cohort data combined with correctly calculated p-values and AIC values).

3.4.1.2.1.1 Reality Check Using Endogenous/Background Level Data

Although USEPA's 2016 dose-response assessment was based on exogenous EtO exposures, by corollary it applies to the corresponding internal doses that produce excess cancer risk. USEPA acknowledges that endogenous doses may contribute to background risk (pp. 4-95 to 4-96 of USEPA 2016), and in fact use their URF to extrapolate to risk-based air concentrations (1E-06 to 1E-04) corresponding to doses well below normal endogenous doses (e.g., the USEPA 1E-04 air concentration corresponds to a dose that is over 30 times lower than even the 1st percentile of normal endogenous in non-smokers). If such miniscule additive doses to endogenous background are associated with excess risk (as predicted by the USEPA URF) despite being subject to the same defense mechanisms as much higher endogenous doses, then it stands to reason that each equivalent endogenous dose subject to the same defense mechanisms would be expected to result in the same risk (i.e., the same risk per unit internal dose). This is consistent with the basic toxicological principle that equal internal doses give rise to equal risk. That is, in dose-response/risk assessment it is standard practice to consider equal internal doses as equipotent in producing carcinogenic effects (e.g., use of PBPK modeling to extrapolate between species and/or different exposure pathways). Thus, the standard risk assessment practice of considering equal internal doses as equipotent in producing carcinogenic effects underlies TCEQ's application of the USEPA URF used to extrapolate to very low internal EtO doses to higher internal endogenous doses of the same chemical.

Consistent with the discussion above, endogenous and background level data can be used for a reality check on the USEPA (2016) lymphoid cancer URF. Use of the EtO air concentration corresponding to the mean of normal endogenous background levels in the unexposed population (1.9 ppb) in conjunction with the USEPA (2016) age-dependent adjustment factor (ADAF)-adjusted URF for lymphoid cancer (7.1E-03 per ppb) suggests a background incidence of ≈1.35% in non-smokers due to endogenous EtO alone, which remarkably would be almost half (46%) of the lymphoid cancer background incidence of 3% (p. 4-95 of USEPA 2016). However,

the smoking population must also be considered. Use of the EtO air concentration corresponding to the mean background in smokers (18.8 ppb at an HEV of 205.4 pmol/g Hb; Tables 2 and 4 of Kirman and Hays 2017) along with that for non-smokers (1.9 ppb) and USEPA's lymphoid cancer URF (4.8E-03 per ppb, ADAF unadjusted) with ADAFs for early-life exposure (at 1.9 ppb) suggests an incidence of lymphoid cancer in smokers of ≈8% due to EtO alone. This estimate for smokers is particularly telling because: (1) The significant (i.e., 10-fold) increase in internal EtO dose is due to exogenous exposure (i.e., smoking), for which the URF inarguably applies; and (2) This URF-predicted incidence would make lymphoid cancer (i.e., leukemia, non-Hodgkin's lymphoma, multiple myeloma) about as common as lung cancer in smokers (e.g., lifetime lung cancer risk for current smokers of ≈8-14%; Bruder et al. 2018), but it is not and the rate ratios (RRs) are much lower (e.g., see lung cancer versus myeloid leukemia results in Figures 3 and 4 of Gandini et al. 2008 and Table 1 of Doll et al. 2005; see lung/bronchus cancer versus lymphoma, multiple myeloma, and leukemia RRs in Jacob et al. 2018), even despite the plethora of additional carcinogens in tobacco smoke.

Weighting the URF-estimated lymphoid cancer incidence for smokers (8%) at above 25% of the population (e.g., for 1980-2005 (Wang et al. 2018) since current cancer rates would reflect contributions from past smoking, consistent with the USEPA 2016 exposure lag period of 15 years) with that for non-smokers (1.35%) results in a population estimate greater than the lymphoid cancer background incidence of 3% cited by USEPA (p. 4-95 of USEPA 2016) due to background EtO levels in the U.S. population alone. That is, without contributions from other potential causes of lymphoid cancer such as known chemical leukemogens, contributions from the endogenous conversion of ethylene to EtO, other risks factors such as genetic predispositions, etc. This reality check, as well as the smoking-specific one, suggests a scientifically unreasonable URF.

3.4.1.2.1.2 Endogenous Conversion of Exogenous Ethylene to EtO: Potential Risk Implications based on USEPA (2016)

In regard to the endogenous conversion of exogenous ethylene to EtO, USEPA may not have adequately explored the potential contributions of ethylene exposure to EtO risk, stating "only ≈3% of exogenous ethylene was converted to EtO in workers exposed to 0.3 ppm" and that "exogenous ethylene exposure is unlikely to contribute significantly to the effects associated with exposure to exogenous EtO in humans" (p. 3-30 of USEPA 2016). *However, based on USEPA's URF, mean environmental concentrations of ethylene in many areas would in fact result in unacceptable excess risk estimates when multiplied by 0.03 to account for the USEPA-cited endogenous conversion of 3% of exogenous ethylene to EtO.* More specifically, greater than a 1E-04 excess risk would be estimated by USEPA's EtO URF at long-term ethylene air concentrations greater than 0.37 ppb (i.e., 0.37 ppb ethylene × 0.03 converted-to-EtO = 0.011 ppb EtO × 9.1E-03 per ppb = 1.0E-04). For additional context, this concentration is not only exceeded by ambient levels in many areas, but also by the indoor mean range (0.82-3.3 ppb)

and personal exposure mean range (3.1-3.6) provided by Health Canada (2016). Interestingly, mean ethylene concentrations reported in human breath (e.g., 23 ppb in Fenske and Paulson 1999, 29-32 ppb reported in Bratu 2019) exceed this 1E-04 excess risk concentration (0.37 ppb ethylene based on USEPA 2016) by over 60-fold. While this example demonstrates what would be important implications of USEPA's EtO URF if it were accurate (see Section 3.4.1.2.2.3), it should not be misconstrued to mean that ethylene realistically represents unacceptable carcinogenic risk (e.g., the International Agency for Research on Cancer (IARC 1994) has classified ethylene as Group 3, not classified as a human carcinogen).

Data on normal endogenous background levels of EtO are also used in the next section, in conjunction with relevant epidemiological data.

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

Key epidemiological findings were reviewed for consistency with expectations for an overall supra-linear dose-response for EtO-induced carcinogenicity. *If the underlying dose-response for EtO-induced cancer in humans were supra-linear with a steep low-dose slope beginning at zero dose, statistically significant increases in critical cancer endpoints would be expected beginning in the lower occupational exposure groups*. That is, if exogenous EtO had a steep dose-response slope (i.e., were a potent carcinogen) starting in the true low-dose region, such as near the range of endogenous levels (as modeled in USEPA 2016), then statistically increased cancer mortality rates would be expected at the "low" worker doses evaluated for large cohorts (NIOSH, UCC), particularly considering that even "low" historical worker exposures have been significantly higher than environmental EtO concentrations (statistically confirmed; see Section 3.4.1.2.2.3). For example, in regard to the NIOSH cohort, Tables IV and V of Hornung et al. (1994) provide measured and estimated worker EtO exposure means (3.5-4.6 ppm) that are ≈1,000,000-2,000,000 times higher than central tendency environmental levels (≈0.0024-0.0034 ppb per USEPA 2016). Remarkably, estimated daily EtO exposure for a job could have ranged from ≈15,000-32,000,000 times the central tendency environmental levels (see Section 3.2.1.1). Mean NIOSH cohort exposure levels of 3.5-4.6 ppm, for example, are over 1,800-2,400 times higher than mean normal endogenous background and about 500-700 times higher than even the 99th percentile of normal endogenous background (Table 4 of Kirman and Hays 2017). For the UCC cohort, the average cumulative EtO exposure was twice as high as that for the NIOSH cohort (although the study power is less; see Section 3.2.1.2). *Thus, considering significantly elevated historical occupational exposures, if EtO-induced cancer had a steep dose-response slope (i.e., were a potent carcinogen) in the true low-dose region (as modeled in USEPA 2016), then the epidemiological evidence for cancer in workers induced by this direct-acting mutagenic carcinogen would be expected to be conclusive, for both males and females, but is not.* For example, regarding the epidemiological evidence, USEPA (2016) states that while there is "*some* evidence of dose-response relationships", "there is *little strength* in the associations."

This certainly would not be expected for a potent low-dose human carcinogen with a steep low-dose slope (as part of the overall supra-linear dose-response) when there was significantly elevated historical occupational exposure, but nevertheless is why USEPA (2016) must partially rely on animal data of dubious human relevance (see Section 3.2.2) for a finding of "carcinogenic to humans." Below, key epidemiological data, alone and in conjunction with data on normal endogenous background levels, are further reviewed to determine if this information is supportive of adoption of a supra-linear dose-response (i.e., the steep low-dose component) for low-dose extrapolation of carcinogenic risk for EtO.

3.4.1.2.2.1 Key Data from the UCC Cohort

Multiple analyses of epidemiological data from the UCC cohort have shown no long-term carcinogenic effects associated with EtO exposure. Swaen et al. (2009) reported no indications of excess cancer risk, including for the lymphohematopoietic malignancies. There were no trends or associations with cumulative exposure for all cause, leukemia, or lymphoid malignancy mortality. Additionally, no statistically significantly elevated SMRs were found in the analysis by hire date, there were no statistically significant increases in the longest duration category, and no suggested trends by duration (all surrogates of exposure). *Likewise, an update of the UCC cohort through 2013 (unpublished as of the date of this DSD) concludes that examination of mortality from all causes of death, all cancers, leukemia, non-Hodgkin's lymphoma, and lymphoid malignancies revealed no evidence for an exposure-related response. EtO exposure in this cohort (with average cumulative dose of ≈67 ppm-years and average follow-up of ≈41 years) was not associated with an observable increase in lymphohematopoietic cancer mortality* (personal communication with Ciriaco Valdez-Flores, an author of a risk assessment paper based in part on the Bender et al. update). These UCC study results in highly-exposed workers are not consistent with EtO-induced carcinogenicity, much less a supra-linear dose-response with a steep low-dose component (e.g., USEPA's linear two-piece spline model).

3.4.1.2.2.2 Key Data from the NIOSH Cohort and Endogenous Data

Regarding key epidemiological data, this section primarily focuses on statistically significant cancer endpoint increases with EtO exposure in the most sensitive sex (male or female) in the NIOSH cohort, although combined results (male + female) are also discussed. Table 5 and Table 6 contain the lowest male or female dose group with a statistically significant increase for each critical cancer endpoint in the cohort based on evaluations by Valdez-Flores et al. (2010) and Steenland et al. (2004, 2003), respectively. Columns 1 and 4 of Table 5 show that based on the analyses in Valdez-Flores et al. (2010), critical cancer endpoints in the NIOSH cohort (i.e., lymphohematopoietic, lymphoid, non-Hodgkin's lymphoma) were only statistically increased in males, and only in the highest (5^{th}) EtO exposure quantile. Breast cancer in females was not statistically increased even in the highest exposure group (5^{th} quantile). The upper ends of the exposure intervals for these highest (5^{th}) quantiles are open ended, and even the lower ends of the exposure intervals are extraordinarily high*. It is remarkable that although workers were*

exposed to EtO air concentrations ≈15,000-32,000,000 times higher than central tendency environmental levels, critical cancer endpoints such as lymphoid cancer mortality were only statistically increased in the highest male exposure group. Such high occupational exposures being required to produce statistically significant increases in a large cohort is not consistent with a steep dose-response slope beginning at zero dose in the low-dose region of a supra-linear dose-response (e.g., in the range of endogenously- or environmentally-relevant doses). Importantly, Table 5 also utilizes data from Kirman and Hays (2017) to calculate environmental EtO exposures (ppm-days) corresponding to the normal endogenous background range (column 2), and then converts those environmental exposures to equivalent occupational exposures (column 3) for comparison to the occupational carcinogenic doses (ppm-days) for critical cancer endpoints (i.e., occupational doses associated with statistically increased cancer). The comparisons provided in column 5 of Table 5 show that across statistically increased cancer endpoints (excludes breast cancer), the lowest carcinogenic doses for EtO (ppm-days, unlagged) in either sex are:

- ≈500-800 times higher than the mean endogenous background EtO dose in the unexposed population;
- ≈600-900 times higher than the median endogenous background EtO dose in the unexposed population;
- ≈1,600-2,700 times higher than the 5th percentile of normal endogenous background EtO doses in the unexposed population; and
- ≈200-300 times higher than even the 95th percentile of normal endogenous background EtO doses in the unexposed population.

The bottom of Table 5 shows that on average, the lower ends of the carcinogenic doses for the most sensitive sex across endpoints are ≈600 higher than the mean endogenous background EtO dose in the unexposed population, ≈700 higher than the median endogenous background EtO dose, ≈2,100 higher than the 5th percentile of normal endogenous background EtO doses, and ≈300 times higher than the 95th percentile of normal endogenous background EtO doses in the unexposed population (Table 5).

Similarly, columns 1 and 4 of Table 6 show that based on the analyses in Steenland et al. (2004, 2003), certain critical cancer endpoints in the NIOSH cohort (i.e., all hematopoietic, lymphoid, non-Hodgkin's lymphoma) were only statistically increased in males, while breast cancer incidence was only statistically increased in females, and only in the highest EtO exposure quantiles for each of these cancer endpoints. The upper ends of the exposure intervals for these highest quantiles are open ended, and even the lower ends of the exposure intervals for these significantly lagged exposures (typically 15 years) are still remarkably high. High occupational EtO exposures being required to produce statistically significant increases in a

large cohort is not consistent with the steep low-dose slope of a supra-linear dose-response starting at zero dose (e.g., across much lower and more environmentally-relevant exposures). The comparisons provided in column 5 of Table 6 show that on average, these lagged carcinogenic doses are:

- ≈90 times higher than the mean endogenous background EtO dose in the unexposed population;
- ≈100 times higher than the median endogenous background EtO dose in the unexposed population;
- ≈300 times higher than the 5th percentile of normal endogenous background EtO doses in the unexposed population; and
- ≈40 times higher than even the 95th percentile of normal endogenous background EtO doses in the unexposed population.

These differences are appreciable considering that the occupational exposures used for these comparisons have been reduced by lagging the exposure 10-15 years, and by using the lowest end of the carcinogenic dose range for each endpoint.

Based on the data in Table 5 and Table 6 (as well as data from the cited source studies), Figure 2, Figure 3, Figure 4, Figure 5, Figure 6, and Figure 7 (Figure 3 and Figure 6 in particular, with a log scale for the x- and y-axis) help demonstrate the significant difference between EtO doses corresponding to the normal endogenous background range (5th-95th percentile) and those associated with (and not associated with) statistically significant increases in the most sensitive sex for critical cancer endpoints in the NIOSH cohort (Valdez-Flores et al. 2010, Steenland et al. 2004, 2003). Figure 4 and Figure 7 help to put into perspective the large differences between occupational EtO doses (ppm-days) and doses corresponding to 1E-06 to 1E-04 excess risk based on USEPA (2016) (i.e., 0.0001-0.01 ppb environmental converted to occupational ppm-days), those corresponding to normal endogenous background levels, and those associated with statistically significant increases in risk*. EtO doses at 1E-06 to 1E-04 excess risk based on USEPA (2016) are orders of magnitude below both those corresponding to normal endogenous background levels and those associated with statistically significant cancer increases in the NIOSH cohort.* Additionally, as shown in Figure 7, although USEPA considers it "*highly plausible* that the dose-response relationship over the endogenous range is *sublinear*," USEPA (2016) applied remarkably steep *supra-linear* model low-dose slopes for lymphoid and breast cancer (see Figures 4-9 and 4-10 of USEPA 2016) in the very region where sublinearity is expected (i.e., ≤ the normal endogenous background range). One consequence is that the EtO air concentration at even the maximum acceptable excess risk (0.01 ppb at 1E-04 risk) is over 30 and 50 times lower than air concentrations corresponding to the 1st and 5th percentiles of normal endogenous background levels, respectively (e.g., 0.56 ppb at the 5th percentile (Table 4

of Kirman and Hays 2017) / 0.01 ppb at 1E-04 risk = 56-fold higher). Put another way, the USEPA considers EtO air concentrations corresponding to more than ≈0.5% percent of mean normal endogenous in non-smokers to be associated with unacceptable risk (i.e., 0.01 ppb/1.9 ppb corresponding to the mean endogenous in non-smokers (Table 4 of Kirman and Hays 2017) × 100 = 0.53%).

In regard to combined (male + female) results, although there were no statistically significant increases in mortality in female workers alone for any critical cancer endpoint in any cumulative EtO exposure group of the NIOSH cohort, combining data from male and female workers results in statistically significant increases for lymphohematopoietic and lymphoid cancer mortality at lower cumulative exposures than when evaluated on a sex-specific basis. For example, although Steenland et al. (2004, 2003) is the source for Table 5, USEPA (2016) had additional analyses conducted for males + females (15-year exposure lag) that showed statistically increased lymphohematopoietic cancer mortality beginning at ≈2,440 ppm-days (midpoint of the 3rd EtO exposure quintile), and increased lymphoid cancer mortality at ≈2,440 ppm-days (midpoint of the 3rd quintile) and ≥13,500 ppm-days (lower end of 5th quintile) (Table 4-2 of USEPA 2016). Similarly, Valdez-Flores et al. (2010) conducted analyses for males + females (no exposure lag) that showed statistically increased lymphoid cancer mortality at ≈2,300 ppm-days (2nd EtO exposure quintile midpoint) and ≥47,559 ppm-days (5th quintile lower end) (Table S.9 of the study). While an explanation of the differences in results for the most sensitive sex versus combined analyses is beyond the scope of this DSD, it is noted that the lower end of these lagged/unlagged carcinogenic EtO doses for NIOSH males + females is ≈50 times higher than lower percentile normal endogenous doses, ≈10-20 times higher than median and upper percentile endogenous doses, and >26,000-37,000 times higher than that associated with typical environmental EtO levels (i.e., background and environmental means of 0.0044-0.0062 µg/m^3 (USEPA 2016) = 0.0024-0.0034 ppb × 70 years × 365 days/year = 61.32-86.87 ppb-days = 0.06132-0.08687 ppm-days).

In summary, high occupational EtO exposure being necessary to produce statistical increases in cancer in the NIOSH cohort, especially in conjunction with null results reported from the UCC cohort (with follow-up through 2013), is not supportive of the steep slope of an overall supra-linear dose-response beginning just above zero dose (e.g., at lower and more environmentally-relevant exposures). *More specifically, risk at endogenous background level doses of EtO (and below) is not expected to be consistent with the lower steep slope portion of an overall supra-linear model (e.g., USEPA's linear two-piece spline model) considering: (1) carcinogenic EtO doses based on the NIOSH study are orders of magnitude higher than normal background endogenous doses in the unexposed population (Table 5* and *Table 6); (2) USEPA's expectation of sublinearity in the endogenous range, which the TCEQ agrees with based on MOA considerations (see Section 3.4.1.1); and (3) the overall study results for both the NIOSH and UCC cohorts in workers exposed to extraordinarily high air concentrations/doses of EtO (e.g.,*

negative findings from the UCC cohort; no statistically increased cancer in the lower NIOSH exposure groups that were still subjected to extraordinarily high air concentrations/doses). In regard to the steep low-dose slope of an overall supra-linear dose-response being inconsistent with cohort findings on EtO-induced carcinogenicity, **the next section demonstrates the statistically significant over-estimation of risk from EtO exposure for the NIOSH cohort by USEPA's selected model assessment for both total lymphoid cancer mortalities in the cohort as well as for every exposure quintile.**

Table 5: Occupational Exposures Corresponding to Normal Background Endogenous Levels of EtO versus Exposures Associated with Statistically Significant Increases in Critical Cancer Endpoints in the NIOSH Cohort (Valdez-Flores et al. 2010 [c])

Statistically Increased Cancer Mortality Endpoint in NIOSH Cohort (sex-specific)	Environmental Exposures Corresponding to Normal Background Endogenous EtO Levels (ppm-days) [a]	Occupational Exposures Equivalent to Environmental Exposures Corresponding to Endogenous EtO Levels (ppm-days) [b]	Occupational Exposure Interval for Lowest Quantile with Statistically Elevated Risk (ppm-days) [c]	Carcinogenic Dose Compared to Normal Endogenous EtO Background Levels
Lymphohematopoietic (statistically increased in males, not females) [d]	48.5 (mean) 40.9 (median) 14.3 (5th percentile) 115.0 (95th percentile)	147.7 124.3 43.5 349.7	≥70,223.59 (highest (5th) quantile)	≥475.6 times ≥564.8 times ≥1,613.6 times ≥200.8 times
Lymphoid Tumors (statistically increased in males, not females) [d]	48.5 (mean) 40.9 (median) 14.3 (5th percentile) 115.0 (95th percentile)	147.7 124.3 43.5 349.7	≥88,348.10 (highest (5th) quantile)	≥598.3 times ≥710.5 times ≥2,030.0 times ≥252.6 times
Non-Hodgkin's Lymphoma (statistically increased in males, not females) [d]	48.5 (mean) 40.9 (median) 14.3 (5th percentile) 115.0 (95th percentile)	147.7 124.3 43.5 349.7	≥117,018.15 (highest (5th) quantile)	≥792.5 times ≥941.1 times ≥2,688.8 times ≥334.6 times
Breast Cancer (*not statistically increased* in females)	48.5 (mean) 40.9 (median) 14.3 (5th percentile) 115.0 (95th percentile)	147.7 124.3 43.5 349.7	≥14,959.26 (highest (5th) quantile not statistically increased; included for context)	≥101.3 times ≥120.3 times ≥343.9 times ≥42.8 times
		Carcinogenic Dose [e] Average Magnitude of Exceedance Over Normal Background Levels at the Endogenous:	**mean median 5th percentile 95th percentile**	**≥622.1 times ≥738.8 times ≥2,110.8 times ≥262.7 times**

[a] Environmental exposure (ppm-days) corresponding to normal endogenous = continuous air concentrations of 0.0019, 0.0016, and 0.00056-0.0045 ppm corresponding to the mean, median, and 5^{th}-95^{th} percentile range for normal endogenous HEV levels in the unexposed (Table 4 of Kirman and Hays 2017) × 70 years × 365 days/year.
[b] Occupational exposure equivalent to environmental (ppm-days) = environmental (ppm-days) × 20 m^3/10 m^3 × 365 days/240 days (i.e., a multiplicative factor of ≈3.042; unrounded values used for calculations); see footnote "2" to Table S.12 of Valdez-Flores et al. (2010).
[c] Only information in the first and fourth columns is based on Table S.9 of Valdez-Flores et al. (2010): Rate ratio analyses and Cox proportional hazards model for cumulative exposure for each combination of endpoint, sex, and study; note that breast cancer was not statistically increased in the rate ratio analysis of Valdez-Flores et al. (2010).
[d] Not statistically elevated in females, only males, so male + female combined results not provided as any risk is driven by the dose-response in males (e.g., statistically significant increases for lymphoid tumors and non-Hodgkin's lymphoma in males + females combined but not in females alone, only males).
[e] These comparisons exclude breast cancer as it was not statistically increased in the rate ratio analyses of Valdez-Flores et al. (2010).

Table 6: Occupational Exposures Corresponding to Normal Background Endogenous Levels of EtO versus Exposures Associated with Statistically Significant Increases in Critical Cancer Endpoints in the NIOSH Cohort (Steenland et al. 2004, 2003 [c])

Statistically Increased Cancer Endpoint in NIOSH Cohort (sex-specific)	Environmental Exposures Corresponding to Normal Background Endogenous EtO Levels (ppm-days) [a]	Occupational Exposures Equivalent to Environmental Exposures Corresponding to Endogenous EtO Levels (ppm-days) [b]	Occupational Exposure Interval for Lowest Quantile with Statistically Elevated Risk (ppm-days)[c]	Carcinogenic Dose Compared to Normal Endogenous EtO Background Levels
All Hematopoietic (statistically increased in males, not females) [d]	48.5 (mean) 40.9 (median) 14.3 (5th percentile) 115.0 (95th percentile)	147.7 124.3 43.5 349.7	≥13,500 (highest (4th) quantile, 15-yr lag)	≥91.4 times ≥108.6 times ≥310.3 times ≥38.6 times
Lymphoid Cell Line Tumors (statistically increased in males, not females) [d]	48.5 (mean) 40.9 (median) 14.3 (5th percentile) 115.0 (95th percentile)	147.7 124.3 43.5 349.7	≥13,500 (highest (4th) quantile, 15-yr lag)	≥91.4 times ≥108.6 times ≥310.3 times ≥38.6 times
Non-Hodgkin's Lymphoma (statistically increased in males, not females) [d]	48.5 (mean) 40.9 (median) 14.3 (5th percentile) 115.0 (95th percentile)	147.7 124.3 43.5 349.7	≥13,500 (highest (4th) quantile, 10-yr lag)	≥91.4 times ≥108.6 times ≥310.3 times ≥38.6 times
Breast Cancer (incidence in females)	48.5 (mean) 40.9 (median) 14.3 (5th percentile) 115.0 (95th percentile)	147.7 124.3 43.5 349.7	>14,620 (highest (5th) quantile, 15-yr lag)	>99.0 times >117.6 times >335.9 times >41.8 times
Carcinogenic Dose Average Magnitude of Exceedance Over Normal Background Levels at the Endogenous:			**mean median 5th percentile 95th percentile**	**≥93.3 times ≥110.9 times ≥316.7 times ≥39.4 times**

[a] Environmental exposure (ppm-days) corresponding to normal endogenous = continuous air concentrations of 0.0019, 0.0016, and 0.00056-0.0045 ppm corresponding to the mean, median, and 5^{th}-95^{th} percentile range for normal endogenous HEV levels in the unexposed (Table 4 of Kirman and Hays 2017) × 70 years × 365 days/year.
[b] Occupational exposure equivalent to environmental (ppm-days) = environmental (ppm-days) × 20 m^3/10 m^3 × 365 days/240 days (i.e., a multiplicative factor of ≈3.042; unrounded values used for calculations); see footnote "2" to Table S.12 of Valdez-Flores et al. (2010).
[c] Only information in the first and fourth columns is based on Tables 4, 6, and 7 of Steenland et al. (2004) and Table 4 of Steenland et al. (2003).
[d] Not statistically elevated in females or females + males, only males.

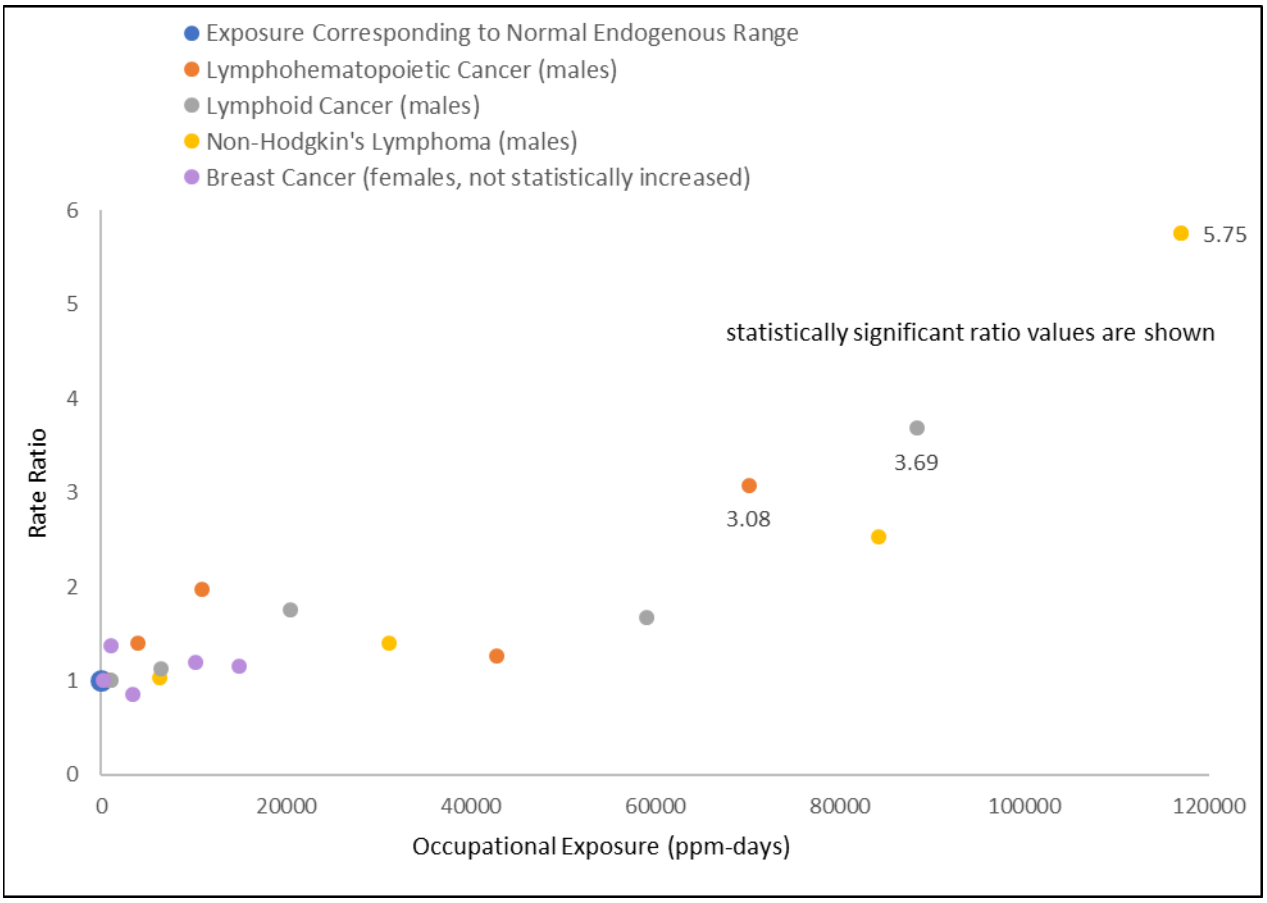


Figure 2: Occupational Exposures Corresponding to Normal Background Endogenous Levels of EtO versus Exposures Associated with Statistically Significant Increases in Critical Cancer Endpoints in the NIOSH Cohort - Linear Scale

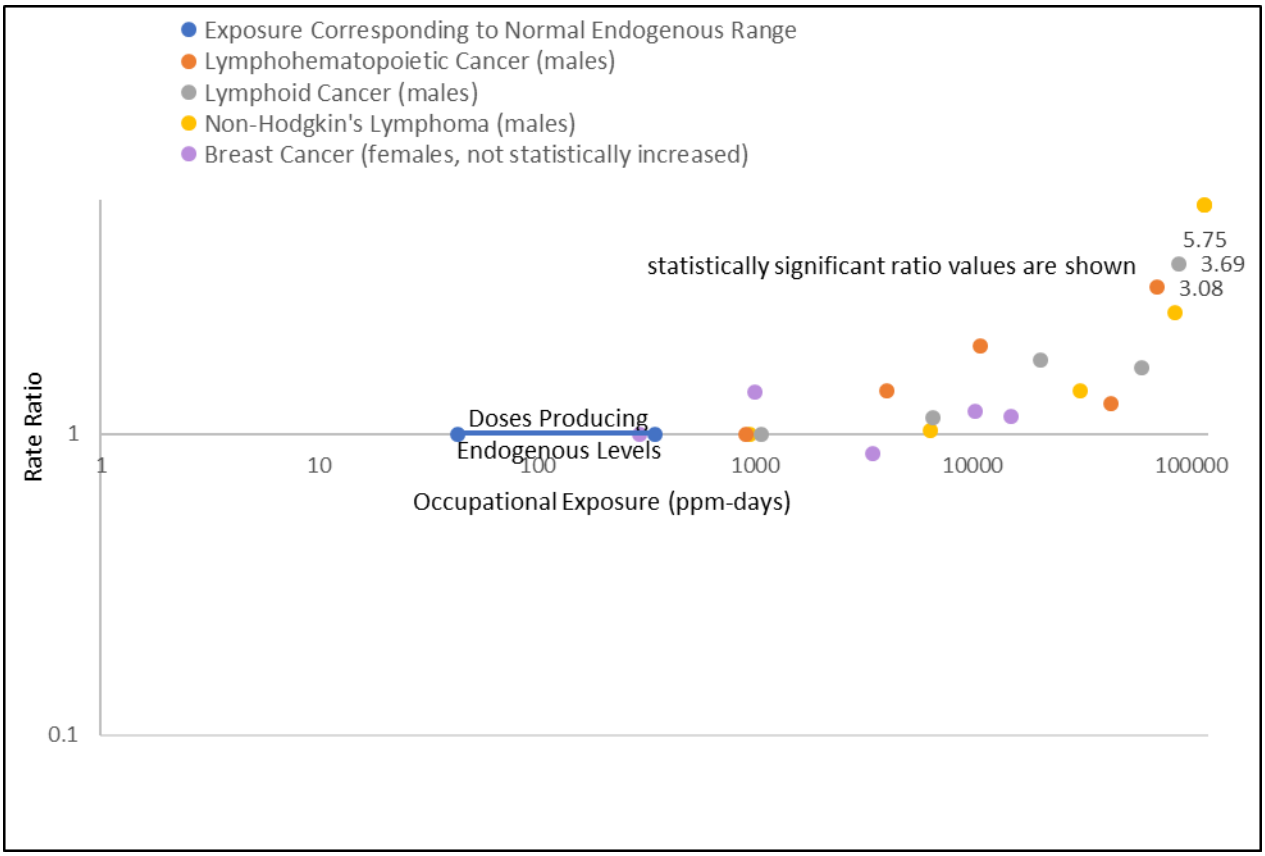


Figure 3: Occupational Exposures Corresponding to Normal Background Endogenous Levels of EtO versus Exposures Associated with Statistically Significant Increases in Critical Cancer Endpoints in the NIOSH Cohort - Log Scale

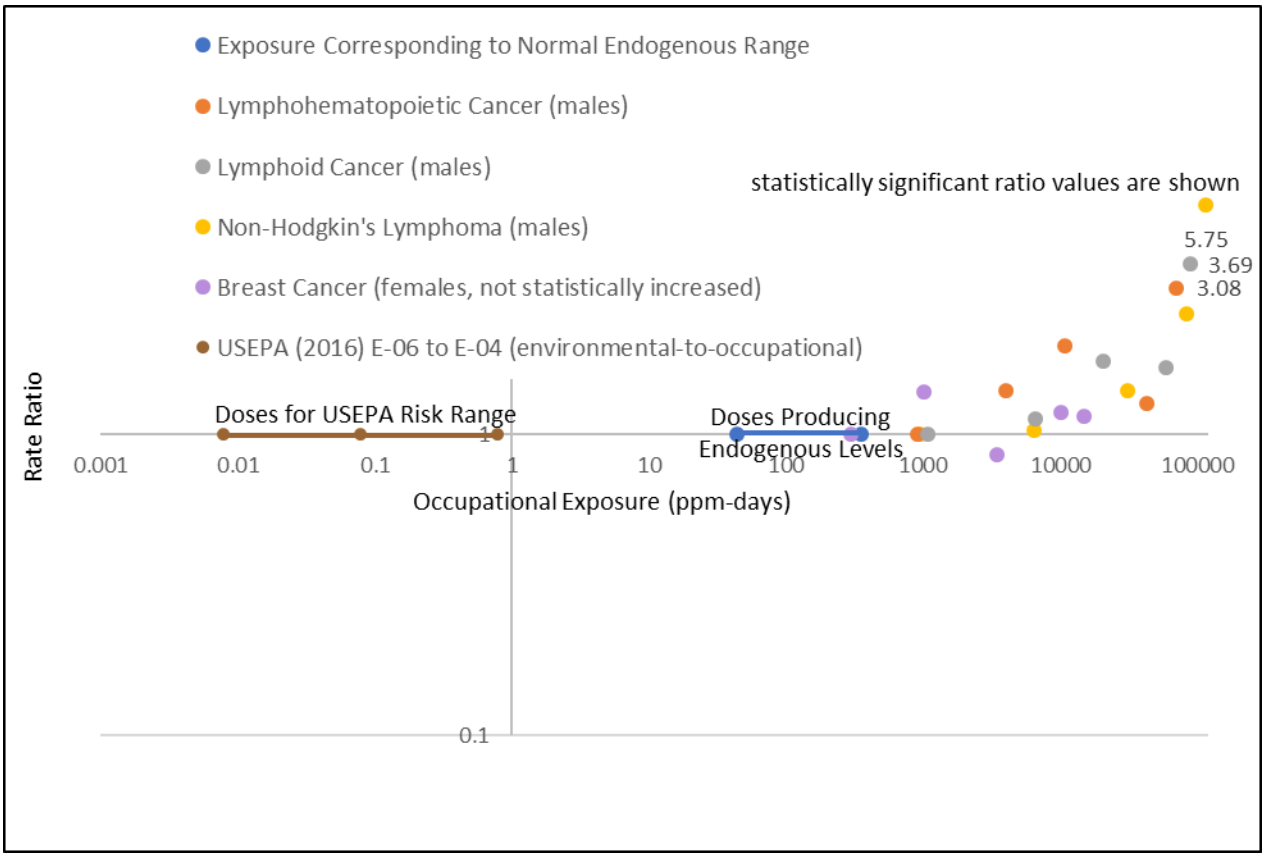


Figure 4: Occupational Exposures Corresponding to USEPA Risk-Based Doses and Normal Background Endogenous Levels of EtO versus Exposures Associated with Statistically Significant Increases in Critical Cancer Endpoints in the NIOSH Cohort - Log Scale

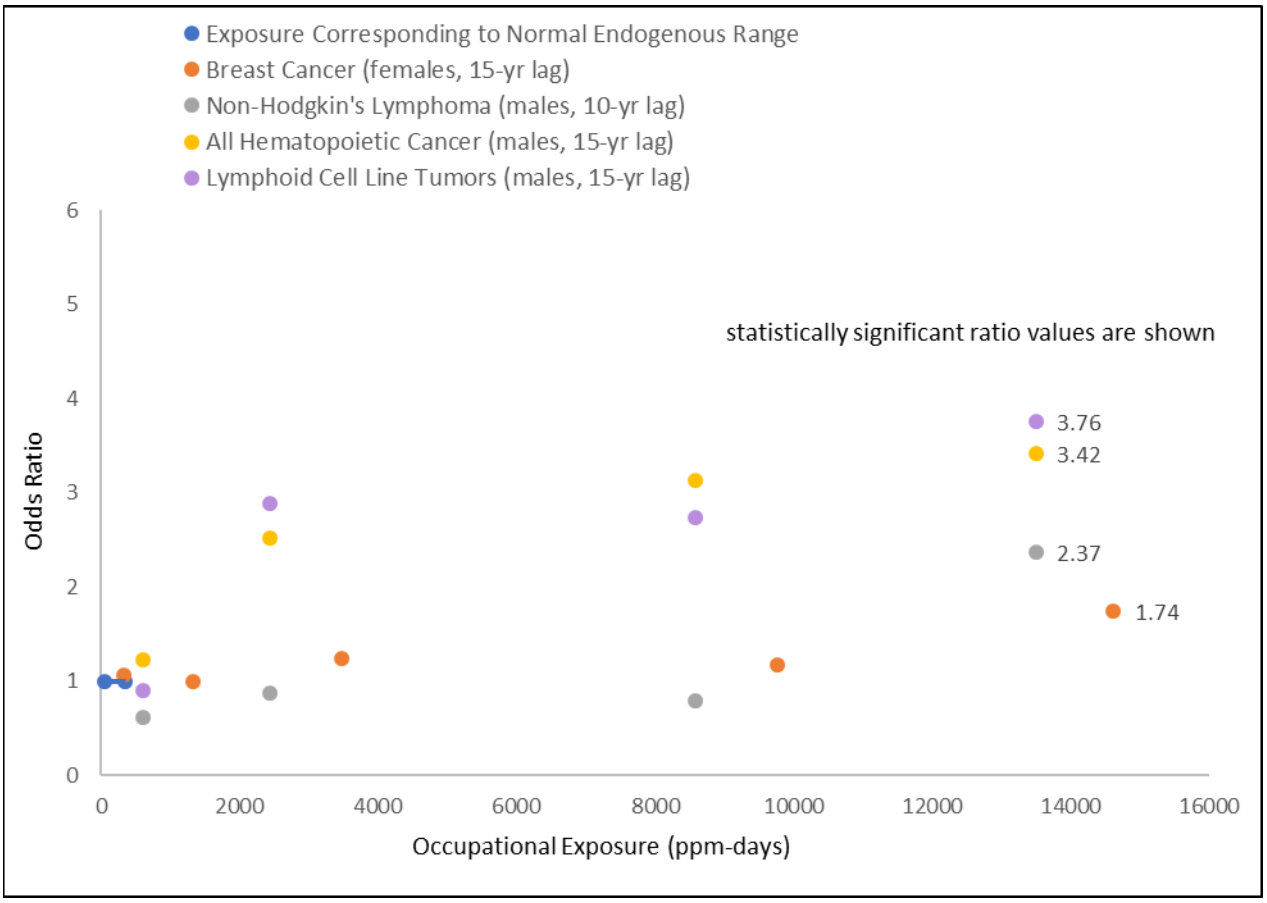


Figure 5: Occupational Exposures Corresponding to Normal Background Endogenous Levels of EtO versus Lagged Exposures Associated with Statistically Significant Increases in Critical Cancer Endpoints in the NIOSH Cohort - Linear Scale

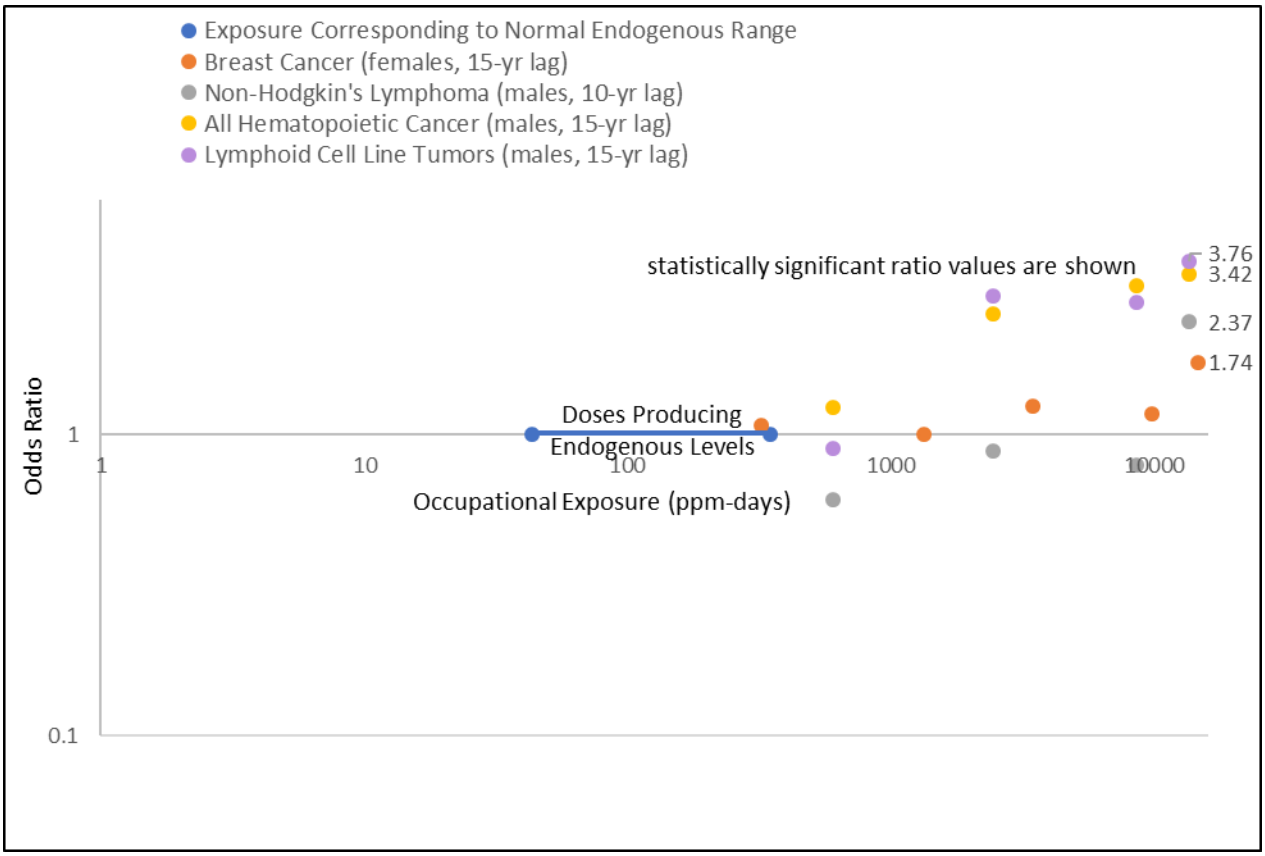


Figure 6: Occupational Exposures Corresponding to Normal Background Endogenous Levels of EtO versus Lagged Exposures Associated with Statistically Significant Increases in Critical Cancer Endpoints in the NIOSH Cohort - Log Scale

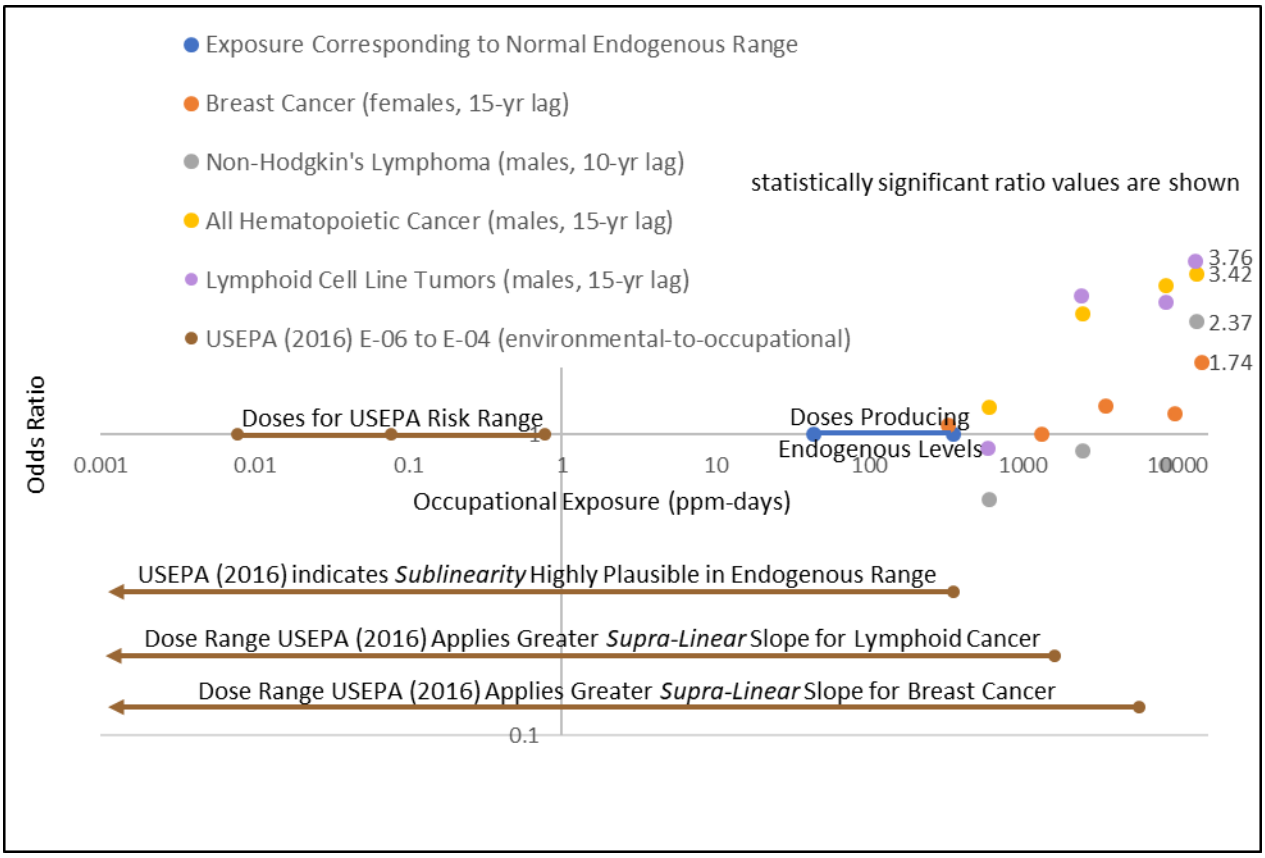


Figure 7: Occupational Exposures Corresponding to USEPA Risk-Based Doses and Normal Background Endogenous Levels of EtO versus Lagged Exposures Associated with Statistically Significant Increases in Critical Cancer Endpoints in the NIOSH Cohort - Log Scale

3.4.1.2.2.3 Lymphoid Cancer in the NIOSH Cohort - Model Predictions Versus Observed

To ground-truth USEPA and other EtO dose-response models (e.g., USEPA's linear two-piece spline model), the various models were used to estimate the number of lymphoid cancer deaths predicted to occur at the EtO exposure levels estimated for the NIOSH cohort compared to the number of cancer deaths that were actually observed in that cohort (details in Appendix 2). As discussed in Section A3.3.1 of Appendix 2, U.S. background hazard rates are appropriate for calculating the model-predicted number of lymphoid cancer deaths due to absence of a healthy worker effect for lymphoid cancer mortality both in the NIOSH cohort specifically and in general. Results demonstrate that *there is no healthy worker effect for this critical endpoint in the key NIOSH worker groups (i.e., male workers who drive lymphoid cancer risk in the cohort, or in male and female workers combined)*. These results based on the NIOSH cohort are consistent with the findings of Kirkeleit et al. (2013).

This model ground-truthing exercise demonstrated that *statistically significant increases in lymphoid cancer mortality would have been observed in every cumulative exposure group beginning in the lowest EtO exposure group of the NIOSH cohort if the model assessment selected by USEPA (i.e., the upper bound of the linear two-piece spline model with the "knot" at 1,600 ppm × days, 15-year exposure lag) were realistic (Table 32 of Appendix 2). In addition, USEPA's selected model assessment predicts that a total of 141 lymphoid cancer deaths (95% CI of 108 to 188) would be expected with the EtO exposure levels estimated for the NIOSH cohort (Table 31 of Appendix 2). However, only 53 total deaths from lymphoid cancers were actually observed, demonstrating that USEPA's selected model assessment statistically significantly over-estimates risk.* By contrast, the model assessment ultimately selected by the TCEQ (i.e., the upper bound on the Cox proportional hazards model, 15-year exposure lag; see Section 3.4.1.4.2) is reasonably accurate, predicting 59 lymphoid cancer mortalities from EtO exposure compared to the 53 actually observed (Figure 8). Although the 95% UCL estimate and not the maximum likelihood estimate (MLE) is being used by USEPA to predict excess risk associated with ambient EtO across the country, the MLE for USEPA's selected model is also statistically significantly over-predictive for the cohort as a whole (Figure 8, Table 31 of Appendix 2). By contrast, the MLE for the model ultimately selected by the TCEQ (i.e., the standard Cox proportional hazards model) is reasonably accurate.

For quintile-specific results (Table 32 of Appendix 2), the model analysis demonstrated that *for every cumulative EtO exposure group, USEPA's selected model assessment (i.e., upper bound on the linear two-piece spline model) statistically significantly over-predicts the 11 lymphoid cancer mortalities that actually occurred in each quintile.* Further, the predictions by the model assessment selected by USEPA (2016) demonstrate that *if the model were realistic, then statistically significant lymphoid cancer increases would have occurred in every cumulative EtO exposure quintile beginning in the lowest (i.e., the lower ends of the 95% CIs range from 17-20 lymphoid cancer mortalities, compared to the 9 lymphoid cancer mortalities in the controls)*.

These predictions by USEPA's selected model assessment are not borne out by the observed cohort data. On the other hand, the log-linear (Cox proportional hazards) model that is ultimately chosen by the TCEQ (see Section 3.4.1.4.2) does not significantly over- or under-predict the lymphoid cancer deaths observed in any NIOSH cumulative EtO exposure group (Figure 9, Figure 10, Figure 11, Figure 12). Similarly, for all but one of the exposure quintiles (quintile 3), the MLE of USEPA's model statistically significantly over-predicts the 11 lymphoid cancer mortalities that actually occurred in each quintile (Figure 9, Figure 10, Figure 11, Figure 12, Table 32 of Appendix 2). By contrast, the MLE for the model ultimately selected by the TCEQ (i.e., the standard Cox proportional hazards model) is reasonably accurate and neither significantly over- or under-predicts the number of lymphoid cancer mortalities.

In summary, as shown here and in Appendix 2, the overall supra-linear model selected by USEPA (2016) over-predicts the number of lymphoid cancer mortalities in the key NIOSH cohort whether based on the assessment selected by USEPA (i.e., 95% UCL for the two-piece spline model) or the associated MLE. That is, *the application of the model assessment selected by USEPA results in statistically erroneous over-predictions of lymphoid cancer risk for the very cohort the model is supposed to fit*. USEPA's selected model assessment is demonstrated to be erroneous for the cohort as a whole and every cumulative exposure group, and the URF derived from it lacks the scientific credibility required for regulatory agency use for this and other reasons described in other sections of this DSD (i.e., the lack of mechanistic justification and other considerations). In contrast to USEPA's over-predictions, the TCEQ's preferred model assessment (i.e., 95% UCL for the standard Cox proportional hazards model) relatively accurately predicts the number of lymphoid cancer mortalities observed in the key cohort. Despite study- and cancer endpoint-specific results that do not support a healthy worker effect for lymphoid cancer, results from a TCEQ sensitivity analysis that nevertheless assumes a health worker effect for lymphoid cancer mortality in NIOSH workers support findings reported in this section (see Section A3.3.2).

Note: USEPA's two-piece linear spline model with the "knot" at 1,600 ppm × days was not the first time their incorrectly calculated model fit criteria (discussed in Section 3.4.1.3) supported a biologically implausible and unpredictive model; *the same model with the "knot" at 100 ppm × days was best supported by their faulty criteria* (Table 4-6 of USEPA 2016). However, USEPA rejected that model as less biologically plausible even in the absence of relevant data, adopting the same model with the "knot" at 1,600 ppm × days as relatively speaking, more biologically plausible/realistic (p. 4-16 of USEPA 2016). While USEPA (2016) utilized no data in making this proclamation, *by contrast, in this DSD the TCEQ utilizes relevant data to put modeling results into a biological plausibility and model predictiveness context* (e.g., endogenous level data, model predictions of the underlying cancer data as a reality check). Had USEPA (2016) used the model best supported by their incorrectly calculated model fit criteria (two-piece linear spline with the "knot" at 100 ppm × days), the statistically significant over-prediction for the cohort

would have been even worse (i.e., MLE estimates of 107.78 lymphoid cancers for the cohort (95% CI of 82.4, 143.9) compared to the 53 actually observed). Either way, the consideration of incorrectly calculated model fit criteria would have led USEPA (2016) to a scientifically unsupportable model. The model ultimately selected by USEPA (with the "knot" at 1,600 ppm × days) is simply the relatively least scientifically unsupportable model of the two two-piece linear spline models considered (i.e., the statistically significant over-predictions would have been even greater for the model USEPA considered best-fitting with the "knot" at 100 ppm × days versus that with a "knot" at 1,600 ppm × days).

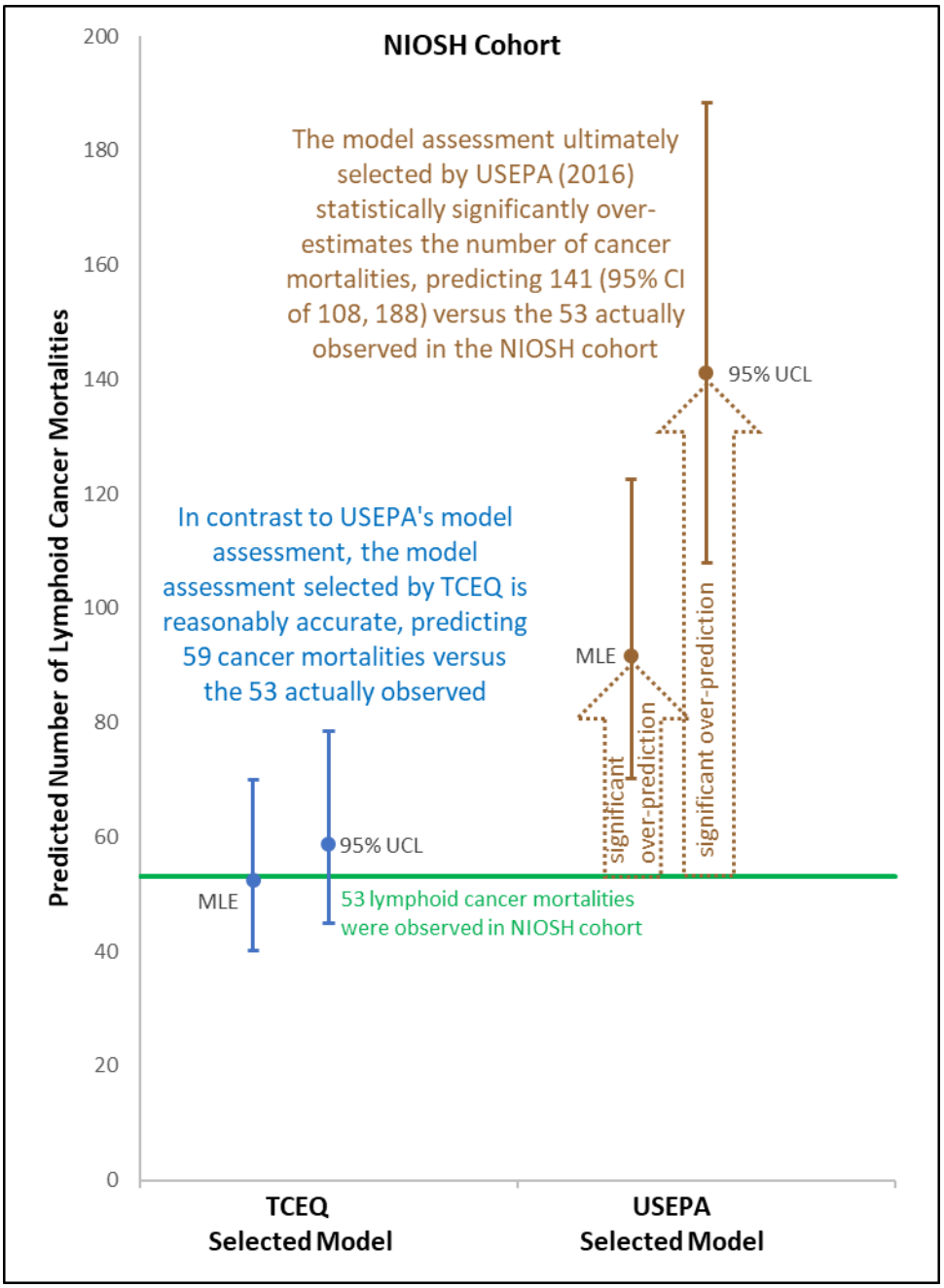


Figure 8: Statistically Significant Over-Prediction of Lymphoid Cancer Mortalities from EtO Exposure by the USEPA (2016) Selected Model Assessment (upper bound of linear two-piece spline) for the NIOSH Cohort versus Reasonably Accurate Results from the TCEQ Selected Model (upper bound Cox proportional hazards)

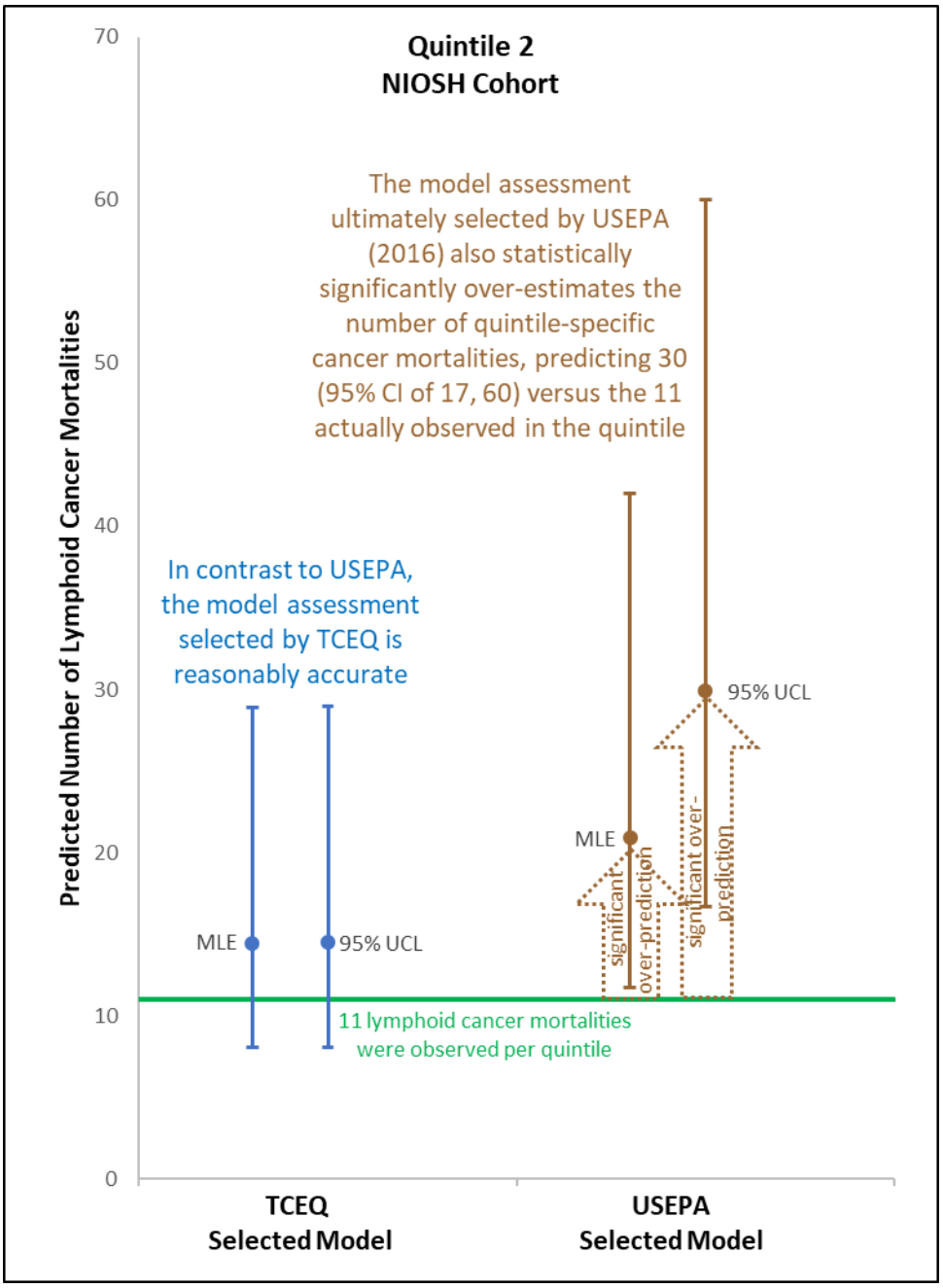


Figure 9: Statistically Significant Over-Prediction of Lymphoid Cancer Mortalities from EtO Exposure by the USEPA (2016) Selected Model Assessment (upper bound of linear two-piece spline) for the NIOSH Cohort versus Reasonably Accurate Results from the TCEQ Selected Model (upper bound Cox proportional hazards) - Quintile 2

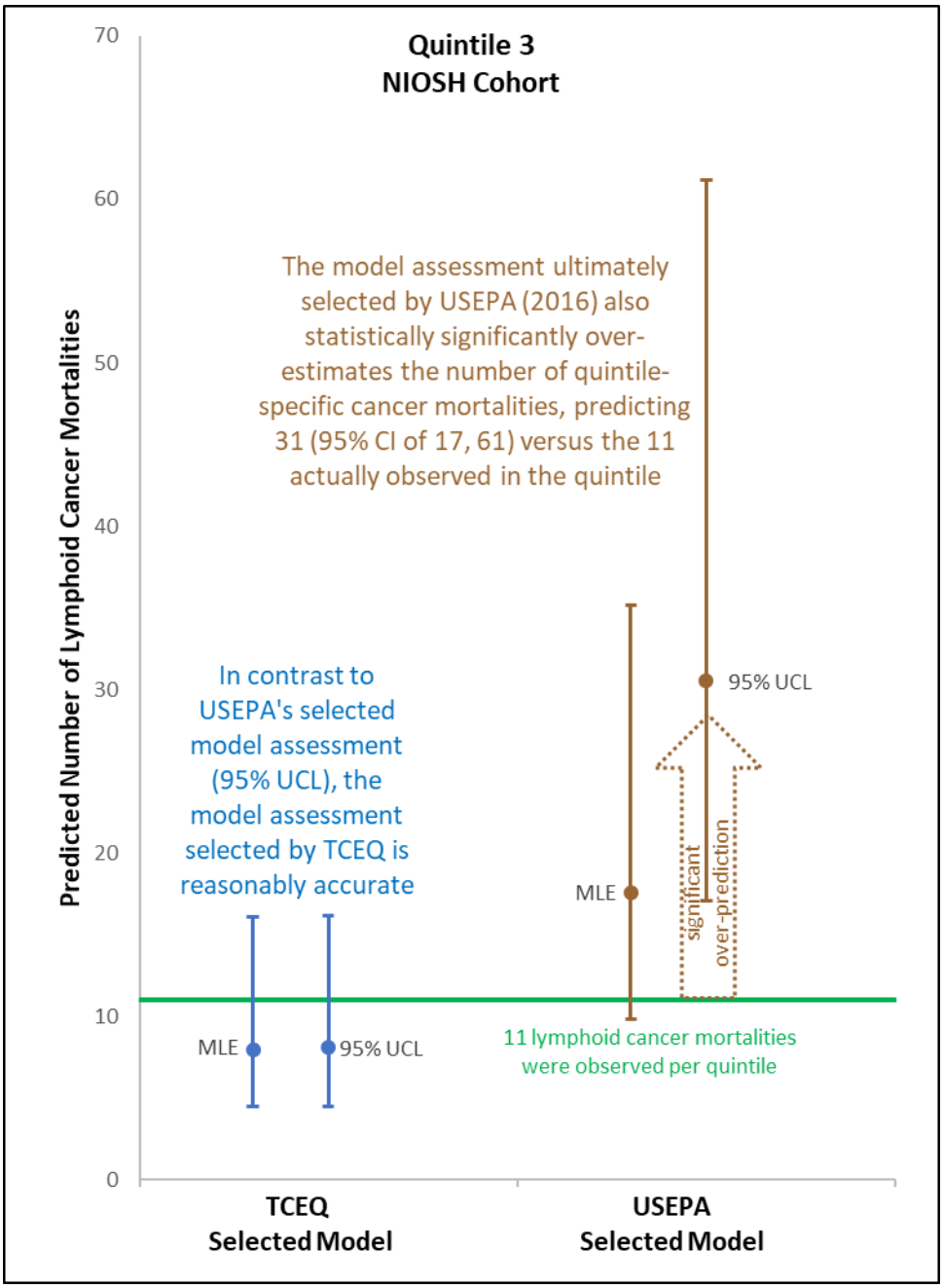


Figure 10: Statistically Significant Over-Prediction of Lymphoid Cancer Mortalities from EtO Exposure by the USEPA (2016) Selected Model Assessment (upper bound of linear two-piece spline) for the NIOSH Cohort versus Reasonably Accurate Results from the TCEQ Selected Model (upper bound Cox proportional hazards) - Quintile 3

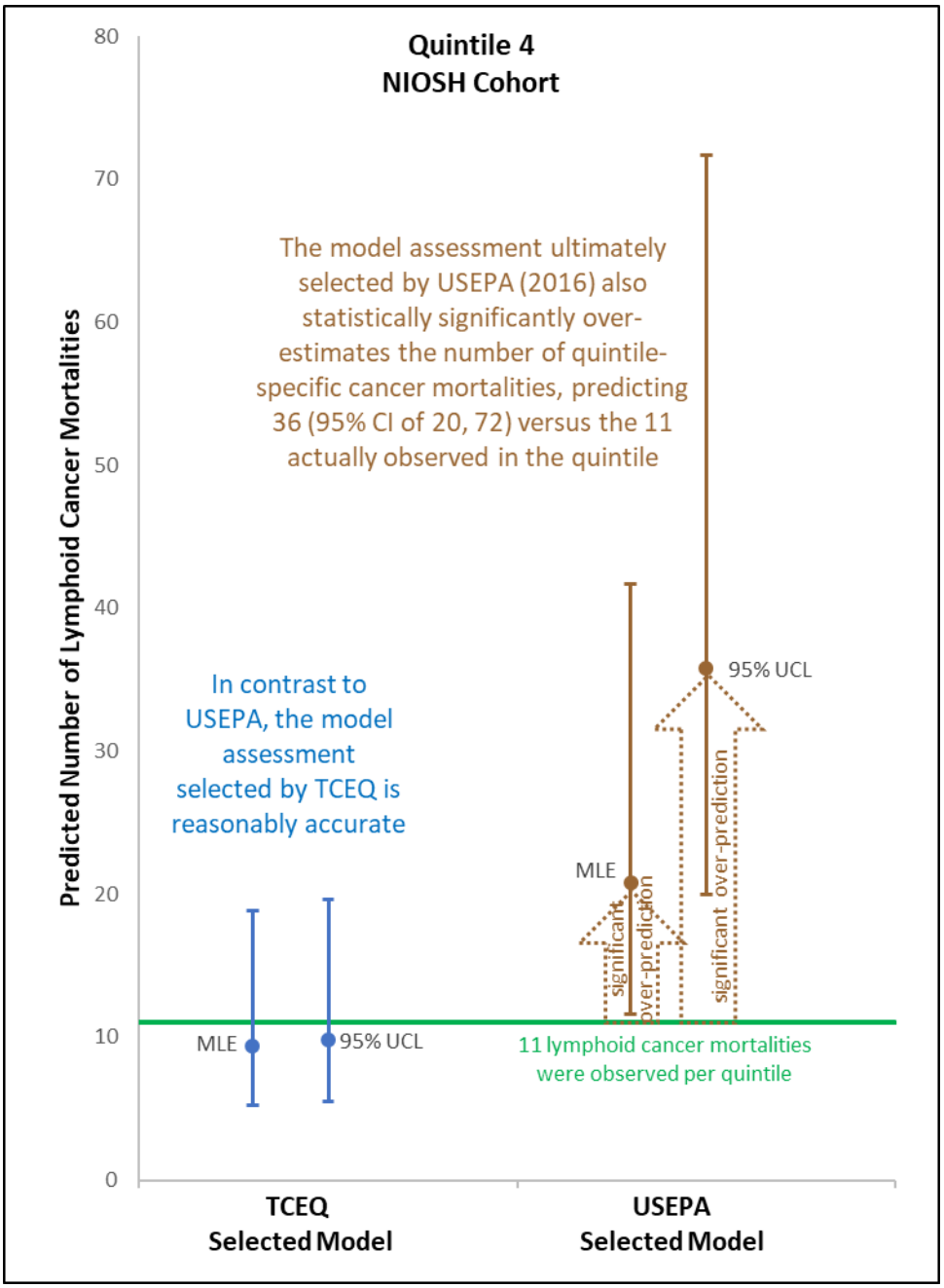


Figure 11: Statistically Significant Over-Prediction of Lymphoid Cancer Mortalities from EtO Exposure by the USEPA (2016) Selected Model Assessment (upper bound of linear two-piece spline) for the NIOSH Cohort versus Reasonably Accurate Results from the TCEQ Selected Model (upper bound Cox proportional hazards) - Quintile 4

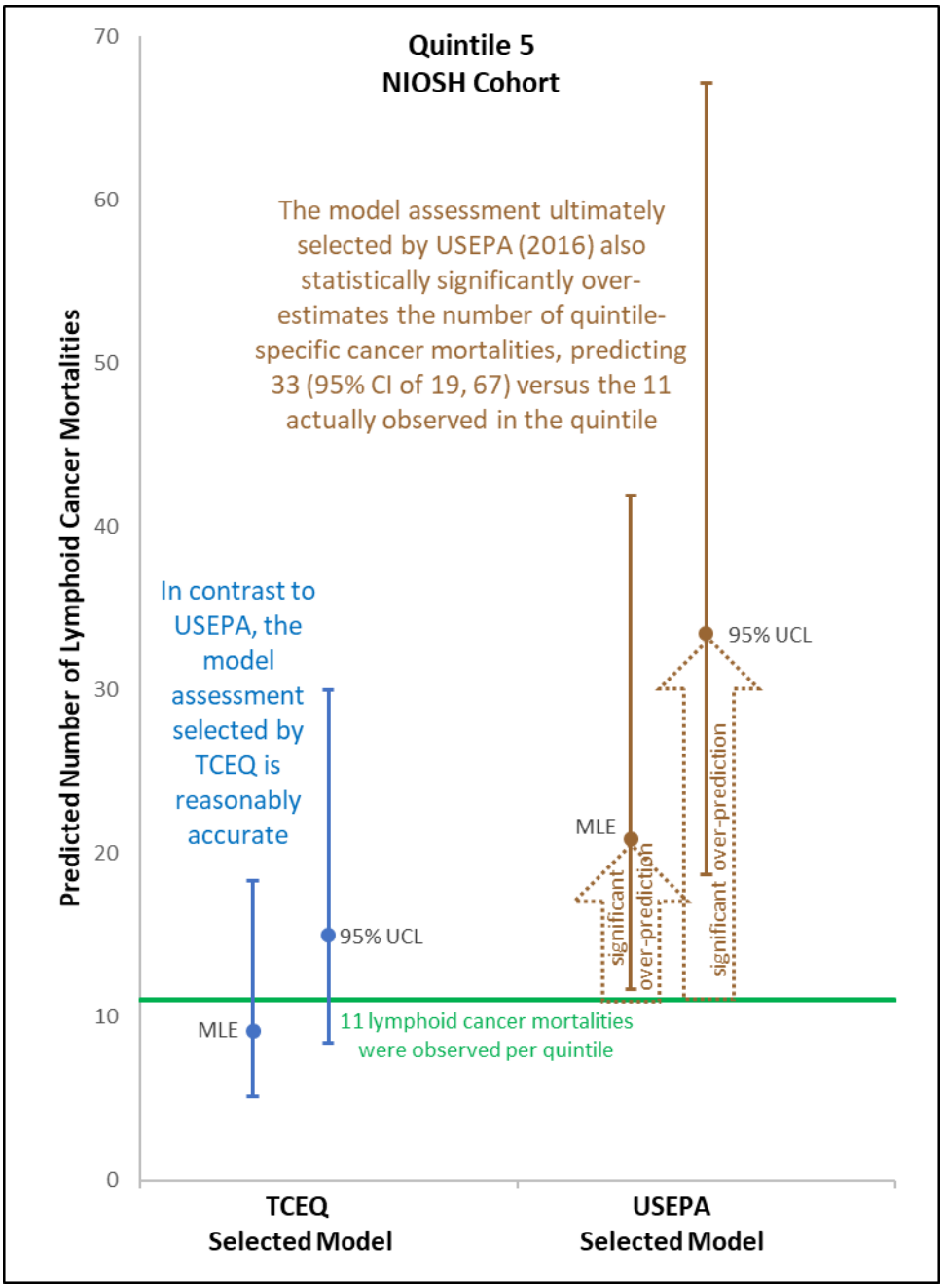


Figure 12: Statistically Significant Over-Prediction of Lymphoid Cancer Mortalities from EtO Exposure by the USEPA (2016) Selected Model Assessment (upper bound of linear two-piece spline) for the NIOSH Cohort versus Reasonably Accurate Results from the TCEQ Selected Model (upper bound Cox proportional hazards) - Quintile 5

3.4.1.2.2.4 Implications of Key Epidemiological Findings, Endogenous Data, and Model Predictions for Use of USEPA's Selected Linear Two-Piece Spline Model Assessment and URF

In summary, the consideration of key epidemiological data such as findings for the UCC and NIOSH cohorts, the carcinogenic doses associated with critical cancer endpoints in the NIOSH study, doses corresponding to normal background endogenous levels, as well as demonstrations of statistically significant over-prediction by USEPA's selected linear two-spline model assessment for lymphoid cancer (the primary URF driver) is not supportive of a supra-linear EtO dose-response (i.e., the steep low-dose slope portion), particularly for low-dose extrapolation (e.g., within or below the range of normal endogenous and ambient levels). Regarding the endogenous range, USEPA (2016) considers it "*highly plausible that the dose-response relationship over the endogenous range is sublinear*." Despite this, as shown in Figure 7, USEPA (2016) actually applied exceptionally steep low-dose slopes from overall supra-linear models for lymphoid and breast cancer in the very low-dose region where a sublinear dose-response is expected (i.e., the endogenous range and even lower). The TCEQ contends that USEPA's choice and application of an overall supra-linear EtO dose-response relationship is therefore internally inconsistent (i.e., self-contradictory). Moreover, the TCEQ has determined that USEPA's use of a supra-linear dose-response (i.e., in particular the steep low-dose slope portion) for low-dose extrapolation is contrary to other considerations discussed above and is not scientifically defensible. *Among other considerations, as part of the scientific weight of evidence, the demonstration of the statistically significant over-estimation of lymphoid cancer risk by the model assessment selected by USEPA (i.e., the upper bound of the linear two-piece spline model with the "knot" at 1,600 ppm × days, 15-year exposure lag), compared to the observed number of lymphoid cancer deaths in the cohort, indicates that this model is not appropriate for deriving the EtO URF.*

[As a peripherally-related topic, the inability to observe sublinearity in the NIOSH cohort might be explained by the lack of dose-response data at low air concentrations (e.g., beginning ≈0.5 ppb) that would allow total internal exposures (endogenous + exogenous) to remain in/near the normal endogenous range. See Figure 3 and Figure 6, keeping in mind that the exogenous exposures corresponding to the normal background endogenous range would themselves produce internal exposures equal to endogenous exposures, over and above them (and that occupational exposures in Figure 6 have been artificially reduced by lagging exposure 10-15 years). Thus, the available dose-response data appear predominated by exposures above the area in the dose-response expected to be sublinear (i.e., within/near/below the normal endogenous range). In such a case, if the available data are at doses sufficiently high to be in the area of the dose-response above the upward inflection point, then the dose-response observed based on the data available might be expected to appear supra-linear overall. Other than providing a hypothetical example in Appendix 3, the TCEQ has not evaluated this possibility further as it is somewhat beyond the scope of this DSD.]

3.4.1.3 Consideration of Model Fit Criteria

Although some models have a biological or mechanistic basis (e.g., Michaelis-Menten model of enzyme kinetics, CIIT biologically-based model for formaldehyde), many models used for dose-response assessment do not (e.g., often only to the extent that low-dose linearity is viewed as consistent with a mutagenic MOA). Thus, in this respect model fit alone is a lesser consideration compared to data (e.g., MOA data) that may (or may not) adequately support use of a particular model (e.g., the overall supra-linear two-piece spline model). For example*, neither USEPA nor TCEQ can cite mechanistic data for EtO that sufficiently support use of a supra-linear model, particularly in the context of concerns about the steepness of the linear two-piece spline model slope over the low-dose region* (e.g., USEPA 2016 considers *sublinearity* "highly plausible" over the endogenous range). In fact, relevant considerations support that the steep low-dose slope of a supra-linear model should not be used over the low-dose region in this case (see Sections 3.4.1.1 and 3.4.1.2). However, model fit is nevertheless a topic of interest for EtO and therefore the topic is discussed, although not as a deterministic consideration on its own when: (1) MOA/mechanistic data must also be considered (TCEQ 2015); and (2) more than one model essentially fits the data equally well but the accuracy of models for predicting the underlying modeled cancer data differs. This section primarily focuses on lymphoid cancer because it was the primary driver of the USEPA (2016) URF, although model fit for breast cancer incidence is also considered. In TCEQ's evaluation of dose-response models that provide the most appropriate fit to the EtO cohort data, the agency also evaluated USEPA's application of model fit criteria to determine if appropriate for use by the TCEQ.

There are two important overarching issues with USEPA's consideration of model fit and ultimate selection of the linear two-piece spline model that the TCEQ must duly consider. The first concerns the statistical optimization of "knot" values for the two-piece spline modeling approach. USEPA (2016) indicates that for this approach, the splines were "fit" to the EtO cancer exposure-response data, and that the knot was generally selected by evaluating different knots in increments (e.g., 100, 500, or 1,000 ppm × days) of cumulative exposure and then by choosing the one that resulted in the best (i.e., largest) model likelihood (pp. 4-13, 4-26, 4-36, and 4-45 of USEPA 2016). Thus, from the process described, it is readily apparent that:

- The "knot" was an iteratively fit model parameter and not simply "preselected" (p. 4-52 of USEPA 2016); and
- The knot values, *being statistically estimated/optimized based on the NIOSH data*, clearly do not conform to the USEPA SAB's notion of potentially fixing some model parameters *not estimated from the data* in the interest of parsimony (see p. 12 of SAB 2015).

"Preselected" is a somewhat ambiguous term that does not adequately characterize and obfuscates how the knot value was statistically fit. This is an important

procedural/methodological issue as it appears that under USEPA's interpretation, multiple model parameters could be statistically estimated/optimized upstream of a final dose-response model, yet none of the fitted parameters would ultimately count as an estimated (*k*) parameter as they were "preselected" based on prior model-fitting exercises. In the present case, the knot values were determined through model fitting with NIOSH data (e.g., maximization of the likelihood of the model for best fit to the lymphoid cancer data), and thus are obviously additional estimated parameters (*k*) in the analysis. That is, for the spline models, the additional parameters (*k*) estimated by USEPA were clearly: (1) the "knot" value; (2) the slope above the knot; and (3) the slope below the knot (*k*=3). However, USEPA (2016) did not account for statistically estimating the optimized knot value. Thus, it appears the degrees of freedom (*df*) were inappropriately reduced for the spline models (i.e., *df=k*, the number of additional parameters estimated for this model over the model with zero-slope with cumulative exposure), which was not inconsequential. Among other consequences, this:

- Inappropriately decreased the p-value for adequate statistical fit, incorrectly implying that the linear two-piece spline model with a knot at 1,600 ppm × days for lymphoid cancer fit the data statistically better than other models in Table 4-6 of USEPA (2016); and
- Inappropriately decreased the Akaike information criterion (AIC) for the spline models, which did not allow for an appropriate comparison of model fit among models for either lymphoid cancer or breast cancer incidence.

Thus, this appears to amount to an unfortunate statistical misevaluation of model fit in USEPA (2016). Appendix D of USEPA (2016), a revised report of Dr. Kyle Steenland submitted in 2010 under contract with USEPA, acknowledges this *df*/p-value issue but then argues for the log-linear two-piece spline model (not ultimately selected by USEPA) not based on statistical fit criteria, but rather conformance with the categorical and cubic spline models in the low-exposure region and the nearly linear exposure-response relationship in that region (p. D-13 of USEPA 2016). In regard to the linear two-piece spline model ultimately selected by USEPA (2016), Section 3.4.1.2.2.3 demonstrates how it statistically significantly over-estimates risk for every cumulative exposure group, including the lowest (see Figure 9, Figure 10, Figure 11, Figure 12).

In regard to the first bullet above, an example in Appendix 4 demonstrates that a p-value of 0.15 is the correct p-value for the likelihood ratio test (not 0.07 as in Table 4-6 of USEPA 2016) when appropriately using *k*=3 for the log-linear two-piece spline model with a knot at 1,600 ppm × days (lymphoid cancer). Similarly, for the linear two-piece spline model with a knot at 1,600 ppm × days (lymphoid cancer) ultimately selected in USEPA (2016), the correct p-value is 0.14 (not 0.07 as in Table 4-6 of USEPA 2016*). Thus, the correct p-values indicate that the likelihoods of these two-piece spline models (linear and log-linear) with knots at 1,600 ppm ×*

days are not different from the likelihood of the null model at the 5% significance level (i.e., the fitted two-piece spline models do not explain the variability in the data statistically significantly better than the null model). *The same is true for all two-piece spline models in Table 4-6 of USEPA (2016) when appropriately using k=3 (p-value range of 0.11-0.15), putting the two-piece spline models and the log-linear (standard Cox regression) model on equal ground in this regard (Appendix 4), although the linear two-piece spline assessment selected in USEPA (2016) statistically significantly over-predicts lymphoid cancer risk whereas the TCEQ's log-linear model does not over- or under-predict risk (see Section 3.4.1.2.2.3, Figure 8, and Appendix 2).*

Regarding the second bullet, the USEPA SAB does not comment on or examine this specific AIC issue in Appendix H of USEPA (2016). *The SAB does recommend less reliance on the AIC* (e.g., pp. I-2 and I-9 of USEPA 2016), *particularly its naïve use without other scientific considerations* (pp. I-17 and I-18 of USEPA 2016), and discusses the true fixing of some model parameters (*as opposed to statistical fitting/estimating parameter values from the data as USEPA did*) in a more general discussion of model parsimony (p. I-16 of USEPA 2016). However, Appendix 4 of this DSD contains an example showing that an AIC of 464.6 is the correct AIC value (not 462.6 as in Table 4-6 of USEPA 2016) when appropriately using *k*=3 for the log-linear two-piece spline model with a knot at 1,600 ppm × days (lymphoid cancer). Similarly, for the linear two-piece spline model with a knot at 1,600 ppm × days (lymphoid cancer) ultimately selected in USEPA (2016), the correct AIC is 464.5 (not 462.1 as in Table 4-6 of USEPA 2016). *Consequently, not only does the linear two-piece spline model for lymphoid cancer ultimately selected by USEPA (2016) not explain the variability in the data statistically significantly better than the null model, but the correct AIC value (464.5) is higher than those for all log-linear (Cox regression) and linear models in Table 4-6 of USEPA (2016). Again, as a related issue, USEPA's selected model assessment for lymphoid cancer mortality statistically significantly over-estimates risk (e.g., Figure 8*, *Figure 9, Figure 10, Figure 11,* Figure 12*).* Appendix 4 also contains an example showing that an AIC of 1,956.360 is the correct AIC value (not 1,954.360 as in Table 4-14 of USEPA 2016) when appropriately using *k*=8 for the linear two-piece spline model with a knot at 5,750 ppm × days (breast cancer incidence) ultimately selected in USEPA (2016). Thus, the correct AIC for the linear two-piece spline model (1,956.360) is higher than or similar to AIC values for the log-linear (Cox regression) and linear models in Table 4-14 of USEPA (2016), with the same being true for the AIC value (1,956.485) for the log-linear two-piece spline model. These two-piece spline model AIC values (for breast cancer incidence) are very similar to the AIC for the standard Cox regression model (1,956.675), which also has a p-value <0.5, putting them on par with each other in this regard (Appendix 4).

As visual fit to the data was also used as a criterion for model selection (e.g., pp. 4-66 and 4-100 of USEPA 2016), the second issue concerns the apparent unintentional visual misrepresentation of model fit in Figures 4-3 and 4-8 of USEPA (2016). Most simply, no true visual comparison of model fit to the data can be made based on USEPA Figures 4-3 and 4-8 (pp. 4-21 and 4-51 of

USEPA 2016) since the data shown are not the data to which the models shown were fit. The actual data underlying model fits shown are the individual data, not the less refined categorical data shown in the figures. *Thus, because the model fits shown in USEPA Figures 4-3 and 4-8 are those to the individual data (and not the categorical data depicted), the figures do not actually show model fit to the modelled data at all*. For lymphoid cancer, objective examination of the model fits to the underlying data reveals no readily apparent superior fit by any particular model (Appendix 5). Unfortunately, the NIOSH breast cancer incidence data are not publicly available, and to the TCEQ's knowledge no graph similar to that in Appendix 5 for lymphoid cancer has been produced to enable an appropriate visual examination of model fits to the actual underlying breast cancer incidence data.

The above statistical and visual fit considerations do not constitute the type of strong data (e.g., robust mechanistic understanding and justification) needed to justify a supra-linear dose-response model (i.e., the steep lower-dose slope component of USEPA's linear two-piece spline model) for low-dose extrapolation for EtO (TCEQ 2015). In summary:

1) Correct p-values for the two-spline models for lymphoid cancer mortality are *within the range* of those for the linear and log-linear (Cox regression) models, although correct AIC values for the two-spline models are *slightly higher* than those for the linear and log-linear (Cox regression) models;
2) *Thus*, even outside of the lack of the most critical deterministic (e.g., mechanistic) data needed to support use of an overall supra-linear model, *there appears to be no strong statistical indication of a need to adopt a non-conventional, supra-linear model over a more standard model*; and
3) As might be expected based on 1 and 2, visual examination of the model fits to the underlying data for lymphoid cancer reveals no readily apparent superior fit by any particular model.

These model fit criteria considerations, especially in conjunction with the consideration of the MOA (Section 3.4.1.1) and model predictions (Section 3.4.1.2.2.3), do not support deviation from more standard/conventional dose-response models (e.g., Cox proportional hazards model).

3.4.1.4 Selection of the Extrapolation Model

3.4.1.4.1 Conclusions on Use of USEPA's Linear Two-Piece Spline Model

The following summarizes the TCEQ's conclusions about the USEPA's linear two-piece spline model for use in the derivation of a URF for EtO. The TCEQ (2015) guidelines require sufficient mechanistic or biological data to support the application of a supra-linear model (i.e., its steep slope beginning at zero dose). *However,* a*doption of an overall supra-linear model (i.e., the*

steep lower-dose component) for EtO low-dose extrapolation is not justified based on mechanistic data or supported by other considerations described above (e.g., key epidemiological data, model ground-truthing). For example, the model assessment ultimately selected by USEPA (i.e., the upper bound of the linear two-piece spline model with the "knot" at 1,600 ppm × days, 15-year exposure lag) has been demonstrated to statistically significantly over-estimate the total number of lymphoid cancer mortalities observed for the NIOSH cohort as a whole (e.g., predicting 141 compared to the 53 actually observed) as well as for every cumulative exposure group (Appendix 2).

The TCEQ's conclusion that relevant considerations do not provide a sufficient scientific basis for the application of an overall supra-linear model (i.e., its steep low-dose slope) for low-dose extrapolation *is consistent with USEPA (2016) acknowledging that reasons (biological, mechanistic, or otherwise) supporting a supra-linear dose-response are unknown, stating to the SAB "the EPA is not aware of a mechanistic explanation" (p. I-29 of USEPA 2016; also see pp. I-34 and 4-71).* In addition to key considerations (i.e., MOA, followed by model predictions of the underlying key cohort data combined with correctly calculated p-values and AIC values), supporting considerations herein suggest that sub-ppb EtO exposure concentrations (e.g., 0.0001-0.01 ppb) may not be consistent with the production of excess (i.e., above background) risk. By contrast, progressively higher EtO air concentrations that produce total internal exposures (endogenous + exogenous) progressively higher than the normal endogenous range are considered more likely to be associated with excess (i.e., above background) risk as the body's normal detoxification and repair processes for endogenous EtO become progressively more likely to be less efficient and/or overwhelmed. For example, a continuous EtO air exposure concentration that itself produces an internal dose above the normal endogenous range is considered most likely to be associated with excess risk (e.g., ≥ ≈4.6-7 ppb; Table 4 of Kirman and Hays 2017), followed by EtO air concentrations that themselves produce internal doses similar to the upper end of the normal endogenous range (e.g., continuous exposure concentrations of 3.5-6.9 ppb would be expected to produce internal doses approximating the 90th to 99th percentile of the normal endogenous range; Table 4 of Kirman and Hays 2017).

USEPA provides no robust biological or mechanistic basis for adopting an overall supra-linear EtO dose-response (i.e., the linear two-piece spline model). In fact, biological/mechanistic considerations by USEPA (2016) are essentially limited to:

1) Choosing between two "knot" values for the two-piece spline models, wherein the agency simply indicates that a knot at…
    a. 1,600 ppm × days (compared to 100 ppm × days) results in a more "biologically realistic" exposure-response for lymphoid cancer as it results in a more gradual rise in low-dose risk and a more plausible rise at higher exposures (p. 4-16 of USEPA 2016), *although USEPA did not actual use any data to put biological*

plausibility into context in dismissing the better-fitting two-piece spline model with the "knot" at 100 ppm × days (compared to the model with the "knot" at 1,600 ppm × days); and

b. 5,750 ppm × days results in a more "biologically realistic" general model shape for breast cancer incidence (p. 4-52 of USEPA 2016).

2) Considering "biologically plausible" exposure lag periods (e.g., pp. D-6 and D-38 of USEPA 2016), although the USEPA SAB did not find USEPA's biological argument to be strong even for this limited purpose (p. I-1 of USEPA 2016).

3) Citing direct mutagenic activity as mechanistic justification for typical default linear low-dose extrapolation (pp. 4-22, 4-37, 4-54, 4-61, 4-74, 4-94, C-30, and I-31 of USEPA 2016).

In acknowledging the lack of mechanistic data for EtO to support the biological plausibility of a supra-linear dose-response, USEPA cites *"insufficient information to elucidate a basis"* (p. I-34 of USEPA 2016). USEPA further indicates that "it is unclear how the available biological data can be used to guide general model selection" (p. I-31 of USEPA 2016). By contrast, the TCEQ utilizes data on endogenous and background EtO levels in combination with key MOA, model prediction, and other data (e.g., see Section 3.4.1.2 and correctly calculated model fit criteria in Appendix 4) to guide and support its model selection as to biological plausibility and the ability to reasonably predict the underlying lymphoid cancer data (as well as not overestimating the general US population background rate). In addition to USEPA's acknowledgment of a lack of a mechanistic and/or biological justification, all the considerations discussed by the TCEQ in various sections above (e.g., MOA, reality checks of model predictions) consistently support the conclusion that there is a lack of data to adequately support the application of the steep low-dose slope of a supra-linear model to extrapolate to significantly lower doses*. The statistical demonstration of the significant over-estimation of lymphoid cancer mortality by the model selected by USEPA (i.e., MLE and upper bound of the linear two-piece spline model with the "knot" at 1,600 ppm × days, 15-year exposure lag) is of particular interest and does not lend scientific credibility to the associated USEPA (2016) EtO URF*.

Without a solid mechanistic basis, USEPA (2016) is primarily left with the "appearance" of supra-linearity based on a less than accurate representation of model fit. The TCEQ considers model fit criteria as a matter secondary to consideration of the most critical deterministic (e.g., mechanistic) data needed to support adoption of a supra-linear model (TCEQ 2015). Regardless, *when appropriately considering statistical and visual model fit, the TCEQ finds no strong (much less compelling) statistical or visual indications of a need to adopt an unconventional, supra-linear model over a more standard model* (see Section 3.4.1.3 and Appendix 5). Moreover, *the underlying key epidemiological data cannot support the application of a supra-linear model (i.e., the steep low-dose slope) for extrapolation to low environmental EtO doses since the data are not informative as to the shape of the dose-response curve at the truly low doses of interest*

(e.g., in the range of typical environmental concentrations). The TCEQ's conclusions about model fit and that the key epidemiological data are not informative as to the shape of the dose-response curve at the low environmental EtO doses of regulatory interest are consistent with USEPA acknowledging that the model and low-dose extrapolation (as well as exposure estimation) are primary sources of uncertainty and *"the actual exposure-response relationship at low exposure levels is unknown"* (pp. 4-61 and 4-74 of USEPA 2016).

High-dose carcinogenicity data alone are incapable of informing truly low-dose risk, no matter how extensive the analyses or peer review (i.e., other relevant information such as mechanism/MOA must be duly considered). *USEPA (2016) should not have based a URF on a supra-linear model (i.e., its lower-dose component) without a robust mechanistic justification for expecting the associated steep low-dose slope component to be applicable at truly low doses or used it to make a large low-dose extrapolation across the endogenous range (and below) considering that the agency actually considers sublinearity as "highly plausible" in this range.* The key study used by USEPA does not provide EtO dose-response data anywhere near environmental levels/doses. USEPA (2005a) recognizes that the relatively small exposure range observed in many epidemiologic studies makes it difficult to discern the shape of the exposure- or dose-response curve, which in the present case concerns the range being limited to only very high exposures. USEPA (2016) acknowledges that *points of departure are substantially above typical EtO environmental levels, resulting in uncertainty in risk at environmental levels* (p. I-14 of USEPA 2016). For example, the agency cites a POD for breast cancer incidence (12 $\mu g/m^3$ or 6.6 ppb) that is almost 3,000 times higher than the cited average background level (0.0044 $\mu g/m^3$ or 0.0024 ppb) and further acknowledges that the two lowest deciles have RRs < 1 and thus "are not by themselves consistent with the unit risk estimate." *Based on all considerations discussed in this DSD (e.g., MOA, normal endogenous levels, model reality checks), the TCEQ finds the assessment to follow to be much more biologically and scientifically reasonable.*

In summary, *robust mechanistic and/or biological data adequate to justify use of an overall supra-linear model (i.e., the application of the steep lower-dose slope for low-dose extrapolation) do not exist for EtO in this case*. In fact, relevant considerations strongly suggest that use of such a model (e.g., USEPA's linear two-piece spline model) for low-dose extrapolation is inappropriate (see the discussions above). As the adoption of supra-linear modeling results (i.e., the steep slope beginning at zero dose for low-dose extrapolation) is scientifically unjustified, the corresponding analyses in USEPA (2016) are considered no further in this DSD for potential adoption by the TCEQ.

3.4.1.4.2 Conclusions on Use of an Alternative Model

Based on the considerations discussed above (e.g., MOA, model ground-truthing), *the TCEQ has determined that a low-dose extrapolation model for EtO carcinogenicity that is no more than linear overall is both reasonable and justified*. The Cox proportional hazards model is one such

model that is linear over the doses of interest, has been used previously by the TCEQ (e.g., in the 1,3-butadiene carcinogenic assessment; TCEQ 2008), and was considered by USEPA (2016). Moreover, Cox regression is the preferred modeling methodology for health endpoints of epidemiology studies under TCEQ guidelines (see Section 7.7.5 of TCEQ 2015). In their assessment of EtO carcinogenicity, Valdez-Flores et al. (2010) also use the Cox proportional hazards model with a default lifetime value of 70 years, consistent with TCEQ guidelines (TCEQ 2015). The European Commission's Scientific Committee on Occupational Exposure Limits has adopted the same modeling approaches of Valdez-Flores et al. for EtO cancer assessment (SCOEL 2012, Valdez-Flores et al. 2011). The standard Cox proportional rate ratio model for EtO includes a single parameter whereas USEPA's two-piece spline models use three parameters. The USEPA SAB indicated that the "AIC can assist with adhering to this principle of parsimony". When correctly evaluated, the AIC for the standard Cox proportional hazards model is somewhat lower than that for USEPA's two-piece spline model, which is preferable (Appendix 4). Thus, use of the standard Cox proportional hazards model abides by the SAB recommendations that "the principle of parsimony (the desire to explain phenomena using fewer parameters) should be considered."

Based on the considerations discussed above, the TCEQ selects the Cox proportional hazards model for the carcinogenicity assessment of EtO. *In summary, use of the standard Cox proportional hazards model is scientifically justified based on*:

1. *MOA* (i.e., the Cox proportional hazards model is indistinguishable from linear across doses of interest and appropriate for dose-response assessment of a direct-acting mutagenic carcinogen, particularly in the acknowledged absence of mechanistic data supporting an overall supra-linear dose-response; see Section 3.4.1.1);
2. *Statistically accurate model predictions of the observed NIOSH lymphoid cancer data* (i.e., the Cox proportional hazards model is shown to neither statistically over- or under-predict the observed data, while USEPA's selected model is demonstrated to be statistically significantly over-predictive; see Section 3.4.1.2.2.3 and Appendix 2);
3. *Reality checks showing that, unlike USEPA's model, it does not over-predict general population background lymphoid cancer risk based on background EtO levels* (e.g., also, if USEPA's assessment were correct lymphoid cancer would be about as common in smokers as lung cancer but is not; see Section 3.4.1.2.1.1);
4. *Biological plausibility when considering endogenous EtO data* (e.g., Cox model risk-based air values correspond to internal doses toward the upper end of the normal endogenous range, consistent with biologically meaningful doses with the potential to produce excess risk distinguishable from background, whereas USEPA's 1 ppt at the same risk level corresponds to an internal dose over 360 times lower than even the 1st percentile of the normal endogenous distribution; see Sections 3.4.1.2.1, 3.4.1.6.2, and 3.4.2); and

5. *Appropriately calculated model fit criteria* (e.g., the more parsimonious Cox proportional hazards model fits the data as well as USEPA's unconventional model and has a slightly lower AIC; see Section 3.4.1.3 and Appendix 4).

Cox proportional hazards modeling results are provided and discussed in the following section.

3.4.1.5 Relevant Cox Proportional Hazards Model Results

In accordance with the section above, Cox proportional hazards modeling results were reviewed. For example, Table 7 provides MLE results from Valdez-Flores et al. (2010) for various potential cancer endpoints.

Table 7: Cancer Endpoint-Specific Environmental EtO Air Concentrations at 1 in 100,000 Excess Risk based on Maximum Likelihood Estimates (MLE) (Valdez-Flores et al. 2010) [a,b,c]

Cancer Endpoint	1E-05 Air Level based on MLE for NIOSH: Males (ppb)	1E-05 Air Level based on MLE for NIOSH + UCC: Males (ppb)	1E-05 Air Level based on MLE for NIOSH: Females (ppb)	1E-05 Air Level based on MLE for NIOSH: Males + Females (ppb)	1E-05 Air Level based on MLE for NIOSH + UCC: Males + Females (ppb)
Lymphoid Tumors [d]	6	10	-ns	8	15
Breast Cancer [d,e]	-ns	-ns	7	17	17
Lymphohematopoietic Tissue [f]	6	10	-ns	9	19
Non-Hodgkin's Lymphoma	12	17	-ns	15	23
Lymphocytic Leukemia	13	16	-ns	19	24
Leukemia	18	23	-SS	78	92
Central Nervous System	-ns	-SS	28	-ns	-ns
Malignant Brain	-ns	-ns	19	-ns	-ns
Pancreatic	-ns	-ns	12	-ns	-ns

[a] Environmental air concentration = occupational concentration × 240 days/365 days × 10 m^3/20 m^3; no occupational exposure lag.
[b] USEPA (2005) age-dependent adjustment factors incorporated.
[c] An EtO air concentration (ppb) value in a cell indicates that the estimated slope was positive for mortality with cumulative ethylene oxide exposure for the cancer endpoint in the NIOSH cohort, though none were statistically significantly positive, while the slopes for other endpoints in the NIOSH cohort were negative (denoted by "-ns") and some even statistically significantly negative (denoted by "-SS").

[d] Cancer endpoint used by USEPA (2016); lymphoid tumors includes non-Hodgkin's lymphoma, multiple myeloma, and lymphocytic leukemia as developed in Steenland et al. (2004).
[e] One male breast cancer mortality in the NIOSH cohort; none in the UCC cohort.
[f] Includes leukemia (and specifically myeloid and lymphocytic leukemia), non-Hodgkin's lymphoma, and multiple myeloma.

Briefly, the Cox proportional hazards model defines a risk set for every case (e.g., every cancer mortality from the specific case), rather than needing a control (i.e., unexposed) group to derive the slope of the relative risk model. The Cox modeling risk sets include all the individuals that are at risk at the time the case occurred (e.g., the time of the cancer mortality from the specific cause); both exposed and unexposed workers. Thus, the TCEQ uses the full risk set, including unexposed and exposed individuals, for every case in the NIOSH study, each possibly having more than 17,000 individuals in the risk set. By contrast, for example, using 100 randomly selected controls for each case (from the pool of all those who survived without the cancer of interest to at least the age of the index case) leads to potentially less precise RRs that are not easily reproducible (e.g., Steenland et al. 2004). This is because of the randomness in the selection of the 100 individuals used compared to using the full risk set for every case.

This DSD considers the same critical cancer endpoints as USEPA (2016), namely lymphoid and breast cancer. However, the results in Table 7 do not incorporate any exposure lag, while USEPA (2016) ultimately utilizes an exposure lag of 15 years. Therefore, in preparing this DSD, the TCEQ contracted with the first author on the Valdez-Flores et al. (2010) study to provide exposure-lagged results that had been previously developed for lymphoid and breast cancer in the course of his research. The TCEQ will ultimately evaluate excess risk results for biological plausibility and scientific reasonableness in the context of relevant information such as normal endogenous EtO levels, associated predicted background rate, etc. Additionally, as referred to in Section 3.3.2, the TCEQ evaluates the weight of evidence for EtO-associated breast cancer in Appendix 7, concluding that USEPA's *carcinogenic to humans* classification is best supported by the lymphoid cancer data (i.e., ultimately, TCEQ's final URF is best based on lymphoid cancer as the critical cancer endpoint).

3.4.1.5.1 Parameter Estimates

3.4.1.5.1.1 Lymphoid Cancer

Table 8, Table 9, and Table 10 contain log-linear (Cox regression) model results for lymphoid cancer mortality in the NIOSH (male + female), NIOSH (male only), and UCC (male only) cohorts, respectively, at various EtO exposure lags. These lymphoid cancer parameter estimates are based on the full NIOSH and UCC datasets (i.e., the individual data and not categorial results). The UCC results are based on an update of the cohort through 2013 that is not yet published.

Table 8: Lymphoid Cell Lineage Tumor Mortality - NIOSH (male + female) - MLE and Standard Error (SE) of the Estimate for Different EtO Exposure Lags

Lag (years)	MLE	(SE)	Deviance [a]: -2 × Ln(Likelihood) (p-value vs null) [b]	Likelihood Ratio Test Statistic: Deviance (null model) – Deviance (model) (p-value vs zero lag) [c]
0	3.48×10^{-6}	(1.83×10^{-6})	726.188 (0.1088)	2.571 (n/a)
5	3.45×10^{-6}	(1.95×10^{-6})	726.495 (0.3224)	2.264 (1.0000)
10	3.11×10^{-6}	(2.23×10^{-6})	727.308 (0.4841)	1.451 (1.0000)
15 [d]	2.81×10^{-6}	(2.65×10^{-6})	727.899 (0.6505)	0.860 (1.0000)
20	1.67×10^{-6}	(3.87×10^{-6})	728.598 (0.9227)	0.161 (1.0000)
25	1.48×10^{-6}	(5.19×10^{-6})	728.687 (0.9646)	0.072 (1.0000)
30	2.03×10^{-6}	(6.74×10^{-6})	728.680 (0.9613)	0.079 (1.0000)

[a] Deviance is -2 × Logarithm of the Likelihood. -2 × Ln (Likelihood) = 728.759 when beta = 0 (null model). The decrease in the deviance at a specific exposure lag (compared with the deviance at 0-years lag) has to be at least 3.84 for the improvement in the deviance to be statistically significant at the 5% significance level. The decrease in the deviance at a non-zero exposure lag (compared with the deviance for the null model) has to be at least 5.99 for the improvement in the deviance to be statistically significant at the 5% significance level.

[b] p-value vs null compares the maximum likelihood of the model fit to the maximum likelihood of the null model. A small p-value indicates that the model with the specified lag fits the data better than the null model.

[c] p-value vs zero lag compares the maximum likelihood of the model fit with the specified lag to the maximum likelihood of the model with zero lag. A small p-value indicates that the model with the specified lag fits the data better than the model with zero lag. None of the exposure lags result in a model that fits the cancer data statistically significantly better than the model with no lag at the 5% significance level.

[d] Exposure lag ultimately used by USEPA (2016).

Table 9: Lymphoid Cell Lineage Tumor Mortality - NIOSH (male only) - MLE and SE of the Estimate for Different EtO Exposure Lags

Lag (years)	MLE	(SE)	Deviance [a]: -2 × Ln(Likelihood) (p-value vs null) [b]	Likelihood Ratio Test Statistic: Deviance (null model) – Deviance (model) (p-value vs zero lag) [c]
0	3.89×10^{-6}	(1.77×10^{-6})	354.312 (0.0696)	3.293 (n/a)
5	3.85×10^{-6}	(1.89×10^{-6})	354.761 (0.2412)	2.844 (1.0000)
10	3.47×10^{-6}	(2.17×10^{-6})	355.795 (0.4045)	1.810 (1.0000)
15 [d]	3.12×10^{-6}	(2.61×10^{-6})	356.553 (0.5910)	1.052 (1.0000)
20	1.63×10^{-6}	(4.08×10^{-6})	357.467 (0.9333)	0.138 (1.0000)
25	6.50×10^{-7}	(6.06×10^{-6})	357.594 (0.9945)	0.011 (1.0000)
30	1.70×10^{-6}	(8.66×10^{-6})	357.604 (0.9995)	0.001 (1.0000)

[a] Deviance is -2 × Logarithm of the Likelihood. -2 × Ln (Likelihood) = 357.605 when beta = 0 (null model). The decrease in the deviance at a specific exposure lag (compared with the deviance at 0-years lag) has to be at least 3.84 for the improvement in the deviance to be statistically significant at the 5% significance level. The decrease in the deviance at a non-zero exposure lag (compared with the deviance for the null model) has to be at least 5.99 for the improvement in the deviance to be statistically significant at the 5% significance level.

[b] p-value vs null compares the maximum likelihood of the model fit to the maximum likelihood of the null model. A small p-value indicates that the model with the specified lag fits the data better than the null model.

[c] p-value vs zero lag compares the maximum likelihood of the model fit with the specified lag to the maximum likelihood of the model with zero lag. A small p-value indicates that the model with the specified lag fits the data better than the model with zero lag. None of the exposure lags result in a model that fits the cancer data statistically significantly better than the model with no lag at the 5% significance level.

[d] Exposure lag ultimately used by USEPA (2016).

Table 10: Lymphoid Cell Lineage Tumor Mortality - UCC/Dow 2013 update (males) - MLE and SE of the Estimate for Different EtO Exposure Lags

Lag (years)	MLE	(SE)	Deviance [a]: -2 × Ln (Likelihood) (p-value vs null) [b]	Likelihood Ratio Test Statistic: Deviance (null model) – Deviance (model) (p-value vs zero lag) [c]
0	-1.42×10^{-5}	(9.17×10^{-6})	299.443 (0.0592)	3.559 (n/a)
5	-1.50×10^{-5}	(9.44×10^{-6})	299.216 (0.1506)	3.786 (0.6338)
10	-1.58×10^{-5}	(9.74×10^{-6})	299.021 (0.1366)	3.981 (0.5159)
15 [d]	-1.60×10^{-5}	(9.94×10^{-6})	299.059 (0.1392)	3.943 (0.5355)
20	-1.52×10^{-5}	(9.91×10^{-6})	299.497 (0.1733)	3.505 (1.0000)
25	-1.53×10^{-5}	(1.03×10^{-5})	299.744 (0.1961)	3.258 (1.0000)
30	-1.51×10^{-5}	(1.07×10^{-5})	300.156 (0.2410)	2.846 (1.0000)

[a] Deviance is -2 × Logarithm of the Likelihood. -2 × Ln (Likelihood) = 303.002 when beta = 0 (null model). The decrease in the deviance at a specific exposure lag (compared with the deviance at 0-years lag) has to be at least 3.84 for the improvement in the deviance to be statistically significant at the 5% significance level. The decrease in the deviance at a non-zero exposure lag (compared with the deviance for the null model) has to be at least 5.99 for the improvement in the deviance to be statistically significant at the 5% significance level.

[b] p-value vs null compares the maximum likelihood of the model fit to the maximum likelihood of the null model. A small p-value indicates that the model with the specified lag fits the data better than the null model.

[c] p-value vs zero lag compares the maximum likelihood of the model fit with the specified lag to the maximum likelihood of the model with zero lag. A small p-value indicates that the model with the specified lag fits the data better than the model with zero lag. None of the exposure lags result in a model that fits the cancer data statistically significantly better than the model with no lag at the 5% significance level.

[d] Exposure lag ultimately used by USEPA (2016).

In regard to Table 8, Table 9, and Table 10, none of the EtO exposure lags result in a model that fits the NIOSH and UCC lymphoid cancer data statistically significantly better than the log-linear (Cox regression) model with no lag (at the 5% significance level). Aside from this statistical consideration, which does not give rise to a preference for any particular exposure lag duration, from a biological perspective it is reasonable to include an exposure lag of some duration to account for a latency period between exposure and cancer. *For this reason, as well as consistency with USEPA (2016), the TCEQ will also utilize an exposure lag of 15 years*.

3.4.1.5.1.2 Breast Cancer

In addition to lymphoid cancer, USEPA (2016) utilizes breast cancer incidence (subcohort with interviews) as a cancer endpoint. Unfortunately, the NIOSH breast cancer incidence data were not publicly available for independent analysis. Therefore, the TCEQ will consider log-linear (standard Cox regression) 15-year exposure-lagged model results for breast cancer incidence (subcohort with interviews) from USEPA (2016). Table 11 contains relevant results adapted from Table 4-12 of USEPA (2016).

Table 11: Breast Cancer Incidence (with interviews) - NIOSH (females) - MLE and SE of the Estimate [a]

Model	Lag (years)	MLE	(SE)
log-linear (standard Cox regression)	15	9.5×10^{-6}	4.1×10^{-6}

[a] Adapted from Table 4-12 of USEPA (2016).

3.4.1.5.2 Risk-Based Air Concentrations and URFs

3.4.1.5.2.1 Lymphoid Cancer

Consistent with the discussions above, 15-year lagged results are highlighted and bolded in Table 12, Table 13, and Table 14 below, which contain environmental EtO air concentrations corresponding to the cited excess risk levels and associated URFs for lymphoid cancer mortality in the NIOSH (male + female), NIOSH (male only), and UCC (male only) cohorts, respectively. The lymphoid cancer calculations include adjustments for ADAFs using the approach described in Sielken and Valdez-Flores (2009). The Cox proportional hazard model was used to directly estimate the 1/100,000 extra risk level, which is at the low end of the observable range, based on the full NIOSH data set (Appendix 7).

Table 12: Lymphoid Cell Lineage Tumor Mortality - NIOSH (male + female) - MLE and 95% Lower Confidence Limit (95% LCL) of the Environmental EtO Concentration at 1 in 100,000 Excess Risk

Lag (years)	MLE Environmental Concentration (1/100,000 excess risk) ppm [a]	95% LCL Environmental Concentration (1/100,000 excess risk) ppm [a]	MLE URF per ppm	95% UCL URF per ppm
0	8.02×10^{-3}	4.30×10^{-3}	1.25×10^{-3}	2.32×10^{-3}
5	8.82×10^{-3}	4.57×10^{-3}	1.13×10^{-3}	2.19×10^{-3}
10	1.08×10^{-2}	4.93×10^{-3}	9.30×10^{-4}	2.03×10^{-3}

Lag (years)	MLE Environmental Concentration (1/100,000 excess risk) ppm [a]	95% LCL Environmental Concentration (1/100,000 excess risk) ppm [a]	MLE URF per ppm	95% UCL URF per ppm
15 [b]	**1.32×10^{-2}**	**5.18×10^{-3}**	**7.57×10^{-4}**	**1.93×10^{-3}**
20	2.49×10^{-2}	5.18×10^{-3}	4.01×10^{-4}	1.93×10^{-3}
25	3.20×10^{-2}	4.73×10^{-3}	3.12×10^{-4}	2.11×10^{-3}
30	2.71×10^{-2}	4.19×10^{-3}	3.69×10^{-4}	2.38×10^{-3}

[a] Environmental concentration = (240 days/365 days) × (10 m^3/20 m^3) × occupational concentration; 1/100,000 excess risk levels were estimated directly from the Cox proportional hazard model, consistent with USEPA (2005a) on selection of a POD at the lower end of the observable range of responses.
[b] Exposure lag ultimately used by USEPA (2016).

Table 13: Lymphoid Cell Lineage Tumor Mortality - NIOSH (male only) - MLE and 95% LCL of the Environmental EtO Concentration at 1 in 100,000 Excess Risk

Lag (years)	MLE Environmental Concentration (1/100,000 excess risk) ppm [a]	95% LCL Environmental Concentration (1/100,000 excess risk) ppm [a]	MLE URF per ppm	95% UCL URF per ppm
0	5.83×10^{-3}	3.34×10^{-3}	1.71×10^{-3}	3.00×10^{-3}
5	6.43×10^{-3}	3.56×10^{-3}	1.56×10^{-3}	2.81×10^{-3}
10	7.84×10^{-3}	3.86×10^{-3}	1.28×10^{-3}	2.59×10^{-3}
15 [b]	**9.67×10^{-3}**	**4.07×10^{-3}**	**1.03×10^{-3}**	**2.46×10^{-3}**
20	2.08×10^{-2}	4.06×10^{-3}	4.81×10^{-4}	2.46×10^{-3}
25	5.94×10^{-2}	3.64×10^{-3}	1.68×10^{-4}	2.75×10^{-3}
30	2.64×10^{-2}	2.81×10^{-3}	3.79×10^{-4}	3.56×10^{-3}

[a] Environmental concentration = (240 days/365 days) × (10 m^3/20 m^3) × occupational concentration; 1/100,000 excess risk levels were estimated directly from the Cox proportional hazard model, consistent with USEPA (2005a) on selection of a POD at the lower end of the observable range of responses.
[b] Exposure lag ultimately used by USEPA (2016).

Table 14: Lymphoid Cell Lineage Tumor Mortality - UCC/DOW 2013 Update (males) - MLE and 95% LCL of the Environmental EtO Concentration at 1 in 100,000 Excess Risk

Lag (years)	MLE Environmental Concentration (1/100,000 excess risk) ppm [a]	95% LCL Environmental Concentration (1/100,000 excess risk) ppm [a]	MLE URF per ppm	95% UCL URF per ppm
0	n/a [c]	2.59×10^{-2}	0	3.86×10^{-4}
5	n/a	4.76×10^{-2}	0	2.10×10^{-4}
10	n/a	1.24×10^{-1}	0	8.06×10^{-5}
15 [b]	**n/a**	**8.70×10^{-2}**	**0**	**1.15×10^{-4}**
20	n/a	3.08×10^{-2}	0	3.25×10^{-4}
25	n/a	2.35×10^{-2}	0	4.25×10^{-4}
30	n/a	1.79×10^{-2}	0	5.58×10^{-4}

[a] Environmental concentration = (240 days/365 days) × (10 m^3/20 m^3) × occupational concentration; 1/100,000 excess risk levels were estimated directly from the Cox proportional hazard model, consistent with USEPA (2005a) on selection of a POD at the lower end of the observable range of responses.
[b] Exposure lag ultimately used by USEPA (2016).
[c] n/a implies that the estimated dose-response relationship was non-increasing.

For lymphoid cancer in the NIOSH cohort (male + female), Table 12 provides an EtO air concentration of 13.2 ppb (1.32E-02 ppm) as corresponding to a no significant excess risk level of 1 in 100,000 based on the MLE for the cohort (15-year exposure lag). Based on the 95% LCL (i.e., LEC_{01}), 5.2 ppb (5.18E-03 ppm) is the EtO air concentration corresponding to a 1 in 100,000 excess risk. *These lymphoid cancer excess risk results are consistent with TCEQ conclusions regarding the range of EtO air concentrations most likely to be associated with excess risk considering normal endogenous levels* (see Sections 3.4.1.2.1 and 3.4.1.4.1). For example, a continuous EtO air concentration ≥ ≈4.6-7 ppb would itself produce an internal exposure above the normal endogenous range and is considered most likely to be associated with excess (i.e., above and distinguishable from background) risk. Thus, the range of 5.2-13.2 ppb for a 1 in 100,000 excess risk is considered biologically plausible considering normal background endogenous levels. Even at a 1 in 1,000,000 excess risk level, the upper end of the corresponding EtO air concentration range (1.32 ppb) would also correspond to a 1 SD increase in internal exposure over normal endogenous levels, which is predicted to be sufficient to move those at the 90th percentile of normal background endogenous levels to over the 95th percentile (Table 4 of Kirman and Hays 2017). This would be a statistically significant increase and could be potentially biologically meaningful/significant (see Section 3.4.1.2.1). Additionally, unlike the

USEPA-selected URF for lymphoid cancer incidence (4.8E-03 per ppb, ADAF unadjusted), use of the URFs corresponding to the MLE (7.57E-07 per ppb) and 95% UCL (1.93E-06 per ppb) along with the EtO air concentrations corresponding to the means of background levels in the unexposed population (1.9 ppb) and smokers (18.8 ppb) and ADAFs for early-life exposure predicts a non-smoker/smoker population-weighted background rate of lymphoid cancer mortality well within the actual background rate. *Accordingly, based on normal endogenous/background levels and this URF reality check, the TCEQ considers these 1 in 100,000 excess risk level EtO air concentrations as both biologically plausible and scientifically reasonable, and therefore has some confidence in them*. Results are similar for NIOSH males only. That is, Table 13 provides MLE and 95% LCL 1 in 100,000 excess risk EtO air concentrations of 9.7 ppb (9.67E-03 ppm) and 4.1 ppb (4.07E-03 ppm), respectively.

For lymphoid cancer in the UCC cohort (males), Table 14 provides an EtO air concentration of 87 ppb (8.70E-02 ppm) that corresponds to a no significant excess risk level of 1 in 100,000 based on the 95% LCL for the cohort (15-year exposure lag), which is 6.6 times higher than the corresponding value based on the NIOSH cohort (males + females). No MLE value is provided because of the negative value in Table 10, consistent with no increased risk with cumulative EtO exposure for the cohort as modeled and reported. The 87 ppb EtO concentration corresponding to a 1 in 100,000 excess risk level for lymphoid cancer (based on the 95% LCL) for the UCC cohort is well above the lower end of the range of continuous EtO air concentrations (≥ ≈4.6-7 ppb) that would itself produce internal exposures above the normal endogenous range, and an EtO air concentration above this range is considered most likely to be associated with excess (i.e., above and distinguishable from background) risk. Even at a 1 in 1,000,000 excess risk level, the corresponding EtO air concentration (8.7 ppb) would also correspond to an increase in internal exposure predicted to be greater than the upper end of the normal endogenous range (Table 4 of Kirman and Hays 2017). Thus, considering normal background endogenous levels, it is biologically plausible that such a change would be associated with excess risk. Additionally, unlike the USEPA-selected URF for lymphoid cancer incidence (4.8E-03 per ppb; ADAF unadjusted), use of the URF corresponding to 95% UCL (1.15E-07 per ppb) along with the EtO air concentrations corresponding to the means of background levels in the unexposed population (1.9 ppb) and smokers (18.8 ppb) and ADAFs for early-life exposure predicts a non-smoker/smoker population-weighted background rate of lymphoid cancer mortality well within the actual background rate. Accordingly, based on normal endogenous/background levels and this URF reality check, the TCEQ considers excess risk at this air concentration as both biologically plausible and scientifically reasonable.

However, the fact that the associated MLE, which represents the best fit to the data (i.e., by definition, the MLE maximizes the likelihood of the observed data), *is consistent with no excess lymphoid cancer mortality risk for the UCC cohort*:

- Suggests that the use of statistical bound results (95% LCL/UCL) for estimating excess risk for both the UCC cohort and other populations (e.g., the general population) may be conservative; and
- Similarly, as part of the weight of evidence, suggests that use of lymphoid cancer excess risk results based on the NIOSH cohort for extrapolation to other populations, even highly exposed occupational populations, may be conservative (especially use of the 95% upper statistical bound on excess risk).

This is further supported by the fact that *none of the slopes for lymphoid mortality in the NIOSH cohort (male + female, male only) or UCC cohort (males) are statistically significantly greater than zero* (at the 5% significance level). Thus, *any excess risk estimates based on these lymphoid cancer analyses may be considered conservative and health-protective*, particularly if the 95% UCL URF is utilized for calculation of the EtO air concentration corresponding to 1 in 100,000 excess risk.

3.4.1.5.2.2 Breast Cancer

Table 15 contains environmental EtO air concentrations corresponding to the cited excess risk level and associated URFs for breast cancer incidence in the NIOSH (female only) cohort.

Table 15: Breast Cancer Incidence (with interviews) - NIOSH (females) - MLE and 95% LCL of the Environmental EtO Concentration at 1 in 100 Excess Risk [a]

Model	Lag (years)	EC_{01} ppm	URF per ppm [b]	LEC_{01} ppm [c]	URF per ppm [b]
log-linear (standard Cox regression)	15	0.126	7.94×10^{-2}	0.0737	1.36×10^{-1}

[a] Adapted from Table 4-15 of USEPA (2016).
[b] URF = 0.01/ EC_{01} or LEC_{01}.
[c] Confidence intervals used in deriving the LEC_{01} was estimated employing the Wald approach for the log-linear RR models.

For breast cancer incidence, the MLE URF (7.94E-05 per ppb) in Table 15 corresponds to an EtO air concentration of 0.13 ppb for a no significant excess risk level of 1 in 100,000 based on females in the NIOSH cohort (15-year exposure lag). Based on the 95% UCL URF (1.36E-04 per ppb), 0.074 ppb is the EtO air concentration corresponding to a 1 in 100,000 excess risk. *These breast cancer incidence excess risk results are not consistent with TCEQ conclusions regarding the range of EtO air concentrations likely to be associated with excess risk considering normal endogenous EtO levels* (see Sections 3.4.1.2.1 and 3.4.1.4.1). For example, the EtO air concentration corresponding to the mean endogenous level in the unexposed population is 1.9 ppb, and continuous exposure to 0.074-0.13 ppb EtO is predicted to correspond to an increase of only ≈6-10% of the SD for normal endogenous levels; a deviation well within the range of

normal endogenous variation in the unexposed population (Table 4 of Kirman and Hays 2017). For additional perspective, the cumulative exposure associated with a lifetime of environmental exposure to 0.074-0.13 ppb (≈1.9-3.3 ppm × days) is >4,430-7,695 times less than the lowest cumulative exposure levels associated with a statistical increase in breast cancer incidence (>14,620 ppm-days, 15-year lagged exposure; Table 5). *The TCEQ finds no basis to conclude that the resulting change in internal dose from continuous exposure to 0.074-0.13 ppb EtO would be biologically meaningful/significant, or from a biological plausibility perspective, would be expected to result in excess risk* (i.e., above and distinguishable from normal endogenous EtO contributions to background risk*).* The TCEQ is not confident in the scientific reasonableness of these excess risk results for breast cancer incidence, particularly given the USEPA-acknowledged insufficient epidemiological weight of evidence for EtO-associated breast cancer (e.g., see Appendix 7).

3.4.1.6 Selection of Critical Cancer Endpoint and URF

3.4.1.6.1 Critical Cancer Endpoint

As discussed above, lymphoid cancer mortality excess risk results are consistent with the range of EtO air concentrations expected to most likely be associated with excess risk. The TCEQ considers the 1 in 100,000 excess risk level EtO air concentrations for lymphoid cancer based on Cox proportional hazards modeling (15-year exposure lag) to be both biologically plausible and scientifically reasonable, thereby increasing confidence in their use for regulatory purposes*. By contrast, the TCEQ lacks confidence in the scientific reasonableness of excess risk results for breast cancer incidence*. For example, exposure to the calculated 1 in 100,000 excess risk EtO air concentrations would result in internal exposures well within the range of normal endogenous variation of non-smokers. *An additional consideration in this lack of confidence is the weak overall weight of evidence for EtO-induced breast cancer in humans (Appendix 7).* USEPA (2016) acknowledges that human data alone are inadequate to classify EtO as *carcinogenic to humans*, and relies on other data, including animal data, to support the insufficient human data (see Section 3.3.2). As discussed in Section 3.2.2.1, the IARC (2019) unanimous consensus is that "At present, the state of the science does not support tumour site concordance as a general principle." *Thus, current best available science indicates that animal data cannot generally be used to support specific sites of chemically-attributable carcinogenesis in humans*. This is even more so the case when laboratory animal results are inconsistent; for example, when EtO induces mammary tumors in mice but not rats. *Accordingly, the laboratory animal data are (to say the least) of dubious relevance for confirmation of, or adequately supporting, the USEPA-acknowledged inadequate epidemiological evidence of breast cancer as a known site of EtO-induced carcinogenesis in humans since*:

> (1) The state of the science does not support tumor site concordance as a general principle (IARC 2019);

(2) Specific to breast cancer, there are significant interspecies differences (i.e., there is relatively little overlap between agents that have been shown to cause breast cancer in humans and animals), with *discordance* generally being the case (IARC 2019) (see Section 3.2.2.1); and

(3) Specific to EtO, animal data are simply not reliable predictors of the sites of EtO-induced carcinogenesis in humans (e.g., lung cancer, brain cancer) (see Section 3.2.2.1).

This is important to realize and acknowledge since *the weight of evidence is that the SIRs/SMRs across individual EtO studies of breast cancer are consistently not statistically significantly elevated*, most being less than 1 (see Table 42 in Appendix 6). It is therefore not surprising that two recent meta-analyses of EtO studies that have examined breast cancer concluded: (1) *With a meta-RR of 0.97 (0.80, 1.18)*, "*Evaluations of workers exposed during sterilization processes do not support the conclusion that EO exposure is associated with an increased risk of breast cancer*" (Marsh et al. 2019); and (2) *"Higher quality epidemiological studies demonstrated no increased risk of breast cancers"* (meta-RR of 0.92 (95% CI: 0.84-1.02), Vincent et al. 2019).

The weight of epidemiological evidence for EtO-induced lympho-hematopoietic cancers may be viewed as somewhat greater. The overall meta-RR for lympho-hematopoietic cancer in Marsh et al. (2019) was 1.48 (1.07, 2.05), although driven by earlier published EtO studies. Additionally, the overall and EtO production/use meta-RRs for Hodgkin's disease were 2.76 (1.21, 6.27) and 5.36 (2.31, 12.44), respectively (Marsh et al. 2019). On the other, in Vincent et al. (2019), only the lympho-hematopoietic cancer meta-RR for low quality studies was statistically significantly greater than 1 (3.55 with a 95% CI of 2.20, 5.75), with reported meta-RRs for high and medium quality studies being 0.98 (0.81, 1.18) and 1.31 (0.83, 2.07), respectively. In regard to specific cancers, although consistency across studies is lacking, statistically elevated risks have been reported for leukemia (SMR of 6.11 (1.7, 15.7)), lymphosarcoma/reticulosarcoma (SMR of 16.93 (3.49, 49.53)), and Hodgkin's lymphoma (SIR of 4.97 (2.38, 9.15)) in certain studies (see p. 5 of Vincent et al. 2019). [USEPA (2016) indicates that EtO causes lymphohematopoietic cancers in both rats and mice, which might be viewed as suggestive of limited additional support since there is 47% overlap between agents that cause lymphoid and haematopoietic cancers in humans and animals (see Section 3.2.2.1), but not necessarily so (IARC 2019).]

Based on the above considerations including the dubious utility of laboratory animal data in supporting the USEPA-acknowledged inadequate epidemiological evidence of EtO-induced breast cancer in humans, the *carcinogenic to humans* classification is best supported by the lymphoid cancer data. *Accordingly, after due consideration of both cancer endpoints, the TCEQ ultimately selects lymphoid cancer as the critical cancer endpoint for derivation of the EtO URF*.

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

For lymphoid tumors, Table 12, Table 13, and Table 14 contain URFs and 1 in 100,000 excess risk EtO air concentrations based on the NIOSH (male + female), NIOSH (male only), and UCC (males) cohorts, respectively. *For protection against lymphoid tumors, a value based on males is considered most conservative.* For example, the URF (MLE) for NIOSH (male + female) is 7.57E-07 per ppb (15-year lag; Table 12) whereas the URF (MLE) for NIOSH (males only) is 1.03E-06 per ppb (15-year lag; Table 13), which is 36% higher. *When determining the final EtO URF, the weighting of data from both cohorts (NIOSH and UCC) must be considered*. For example, in TCEQ's (2011) assessment of the carcinogenicity for nickel a weighting factor of person-years × $1/SE^2$ was used to combine URFs. Similarly, in the carcinogenic assessment of inorganic arsenic (TCEQ 2012), the inverse of the variance ($1/SE^2$) for the β (MLE) was used to weight URFs for the final URF. Inverse-variance weighting (without a person-years weighting factor) is a more standard statistical procedure used in meta-analyses.

SE values for the slopes were obtained from Table 9 and Table 10 (15-year lag) for the Cox proportional hazards model evaluation of lymphoid tumors in NIOSH males (SE=2.61E-06) and UCC males (SE=9.94E-06), respectively. For comparison, it is noted that the SE (2.65E-06;

Table 8) for the full NIOSH cohort (male + female) provides similar weighting results. Both types of weighting factors previously used by the TCEQ were calculated (i.e., $1/SE^2$ and person-years × $1/SE^2$) and are provided in Table 16.

Table 16: Weighting Factors for the Lymphoid Tumor Analyses for the NIOSH and UCC Cohorts

Cohort	Gender	Slope SE	Weight $1/SE^2$	Weight Ratio NIOSH/ UCC	Person-Years	Total Weight Person-Years × $1/SE^2$	Relative Total Weight NIOSH/ UCC
NIOSH	M	2.61E-06	1.47E+11	14.5	189,868	2.79E+16	33.0
NIOSH	M/F	2.65E-06	1.42E+11	14.1	450,906	6.42E+16	76.0
UCC	M	9.94E-06	1.01E+10		83,524	8.45E+14	

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

In accordance with the considerations discussed above, the final EtO URF for lymphoid cancer will be based on the NIOSH (males only) cohort (15-year lagged exposure*)*. Again, modeling results indicate that a lymphoid cancer URF value based on males is conservative for application to females (i.e., results in higher excess risk estimates for females compared to URFs based on males + females combined). *Furthermore, as both a scientifically reasonable and conservative selection, the URF (95% UCL) of 2.5E-06 per ppb will serve as the final URF (ADAF unadjusted) for lymphoid tumors (Table 13).*

EtO Lymphoid Cancer URF = 2.5E-06 per ppb or 1.4E-06 per µg/m^3 (ADAF unadjusted)

The corresponding 1 in 100,000 excess risk EtO air concentration for lymphoid tumors based on this ADAF-unadjusted URF is 4.0 ppb or 7.1 µg/m^3 (i.e., 1E-05/2.5E-06 per ppb = 4.0 ppb; 1E-05/1.4E-06 per µg/m^3 = 7.1 µg/m^3). See Section 3.4.2 for a discussion of the application of ADAFs under USEPA (2005b). A lymphoid cancer 1 in 100,000 excess risk EtO air concentration value based on the full NIOSH (male + female) cohort would be somewhat higher at 5.2 ppb, but within a factor of 1.3. Similarly, based on the URF (MLE) values, EtO air concentrations corresponding to 1 in 100,000 excess risk for both the NIOSH (male + female) full cohort and NIOSH (males only) cohort would be somewhat higher at 13.2 ppb and 9.7 ppb, respectively (Table 12 and Table 13). As stated previously, EtO air concentrations that themselves produce internal exposures above the normal endogenous range (e.g., continuous air exposure concentrations ≥ ≈4.6-7 ppb) are considered most likely to be associated with excess (i.e., above background) risk, and the TCEQ notes that these calculated risk-based air concentrations for lymphoid tumors are remarkably consistent with this expectation.

For additional context, continuous exposure to 4.0 ppb EtO would be predicted to result in an HEV burden (as a biomarker of internal exposure) of approximately 43.6 pmol/g Hb. This HEV level roughly approximates the mean + 1.5 SD (21.1 + 21.9 pmol/g Hb = 43 pmol/g Hb) of the normal distribution in the non-smoking population that results from endogenous EtO exposure (Table 4 of Kirman and Hays 2017). An additional ≈43.6 pmol/g Hb due to continuous exogenous exposure to 4.0 ppb would be predicted to:

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

An exogenous EtO exposure concentration that results in endogenous levels rising above the normal background range in some appreciable portion of the population (e.g., the 90th percentile to > 99th percentile) is considered consistent with the assessment of “excess” (i.e., above background) risk. By contrast, continuous exposure to 0.001 ppb EtO (i.e., the 1 in 100,000 excess risk air concentration using USEPA’s URF) would result in ≈0.0109 pmol/g Hb

added HEV, a mere 0.075% of the SD for normal background endogenous levels and over 360 times less than even the 1st percentile of normal background endogenous levels - this magnitude of change in HEV may be reasonably characterized as biologically insignificant. Extremely low air concentrations corresponding to internal exposure increases that represent such minuscule, *de minimus* fractions of normal endogenous background levels do not provide a scientific basis, much less a robust one, for the biological or mechanistic plausibility of any appreciable excess (i.e., above background) risk. In fact, it suggests the opposite due to the body's inherent ability to deal with typical endogenous levels through normal detoxification and repair processes.

Thus, based on the data evaluated and considerations discussed, an EtO air concentration of 4.0 ppb based on the ADAF-unadjusted URF is considered relatively consistent with the assessment of excess risk as being above and distinguishable from normal endogenous EtO contributions to background risk. In the next section ADAFs are applied to derive the more conservative ADAF-adjusted EtO URF and $^{chronic}ESL_{nonthreshold(c)}$.

3.4.2 Evaluating Susceptibility from Early-Life Exposures

Per Section 3.3.1, the weight of evidence supports a mutagenic MOA for EtO carcinogenicity. The mutagenic MOA is considered relevant to all populations and life stages. See Section 3.5.2 of USEPA (2016) for available information on potentially susceptible life stages and populations (e.g., those with higher HEV adduct levels due to a null GSTT1 genotype or with DNA repair deficiencies). USEPA (2016) indicates that there are no data on the relative susceptibility of children to EtO (e.g., the potential for decreased detoxification/clearance by hydrolysis as a primary metabolic pathway and/or glutathione conjugation). In the absence of chemical-specific data to evaluate potential child/adult differences in susceptibility, USEPA (2005b) provides default ADAFs to account for potentially increased susceptibility in children due to early-life exposure when a chemical has been identified as acting through a mutagenic MOA. Therefore, because of the weight of evidence supporting a mutagenic MOA and the lack of chemical-specific data on potential differences in susceptibility, increased early-life susceptibility should be assumed and ADAFs applied (TCEQ 2015). As previously mentioned, the results utilized by the TCEQ (e.g., Table 13) incorporate USEPA (2005b) ADAFs through the approach described in Sielken and Valdez-Flores (2009). However, USEPA (2016) indicated that this publication misinterprets the application of ADAFs. As such, the TCEQ calculated the ADAF-adjusted $^{chronic}ESL_{nonthreshold(c)}$ for EtO consistent with equation 5-17 of the TCEQ guidelines (TCEQ 2015):

$^{chronic}ESL_{nonthreshold(c)}$ = 6.0E-06/URF = 6.0E-06/2.5E-06 per ppb or 1.4E-06 per µg/m^3 = 2.4 ppb or 4.3 µg/m^3 (ADAF adjusted)

Rounded to two significant figures, the ADAF-adjusted EtO $^{chronic}ESL_{nonthreshold(c)}$ is 2.4 ppb or 4.3 µg/m^3. Note that this value would be insensitive to an additional 10-fold ADAF for *in utero* exposure during the third trimester (i.e., equation 5-17 of TCEQ 2015 would become 5.9E-06/URF = 5.9E-06/2.5E-06 per ppb = 2.36 ppb**). The ADAF-adjusted URF is 4.1E-06 per ppb or 2.3E-06 per µg/m^3** (i.e., 2.5E-06 per ppb × 1.657 (based on equation 5-16 of TCEQ 2015) = 4.1E-06 per ppb).

For additional context, continuous exposure to 2.4 ppb EtO would be predicted to result in an HEV burden (as a biomarker of internal exposure) of approximately 26.2 pmol/g Hb. This HEV level roughly approximates the 75th percentile (26.4 pmol/g Hb) of the normal distribution in the non-smoking population that results from endogenous EtO exposure (Table 4 of Kirman and Hays 2017). An additional ≈26.2 pmol/g Hb due to continuous exogenous exposure to 2.4 ppb would be predicted to:

- Increase the HEV level for the median non-smoker (17.3 pmol/g Hb) to above the 90th percentile (38.8 pmol/g Hb) of normal endogenous background levels; and
- Increase the HEV level in 95th percentile non-smokers (48.7 pmol/g Hb) to the 99th percentile (74.9 pmol/g Hb).

Thus, based on the data evaluated and considerations discussed, an **ADAF-adjusted EtO $^{chronic}ESL_{nonthreshold(c)}$ of 2.4 ppb** is considered health-protective, scientifically reasonable, and relatively consistent with the assessment of excess risk as being above and distinguishable from normal endogenous EtO contributions to background risk. Continuous exposure to 2.4 ppb EtO would be predicted to correspond to internal exposure level increases to the upper end of the range of normal endogenous levels at least for some percentage of the population (e.g., moving those at the 95th percentile to the 99th percentile). Additionally, as shown in Appendix 2, the Cox proportional hazards model assessment used by the TCEQ (log-linear, 15-year exposure lag, 95% UCL) neither statistically over- or under-predicts the lymphoid cancer numbers observed in the NIOSH cohort, but rather is relatively accurate. The calculated EtO $^{chronic}ESL_{nonthreshold(c)}$ (2.4 ppb) falls well within the range (0.13-6.9 ppb) supported by the approach in Kirman and Hays (2017) as protective of human health. Additionally, it is conservative (≈76-92% lower) compared to that proposed by Valdez-Flores et al. (2010) at the 1 in 100,000 excess risk level (i.e., 1-3 ppb at 1E-06 ≈ 10-30 ppb at 1E-05 compared to 2.4 ppb).

3.4.3 Final EtO URF and $^{chronic}ESL_{nonthreshold(c)}$

The ADAF-unadjusted URF is 1.4E-06 per µg/m^3 (2.5E-06 per ppb) based on lymphoid cancer. **The corresponding ADAF-adjusted URF is 2.3E-06 per µg/m^3 (4.1E-06 per ppb)**. **The ADAF-adjusted EtO $^{chronic}ESL_{nonthreshold(c)}$ is 4.3 µg/m^3 or 2.4 ppb,** rounded to two significant figures (see Section 3.4.2).

3.5 Long-Term ESL and Value for Air Monitoring Evaluation

The chronic evaluation resulted in the derivation of the following values for EtO:

- ADAF-unadjusted URF = 1.4E-06 per µg/m^3 (2.5E-06 per ppb) for lymphoid cancer
- ADAF-adjusted URF = 2.3E-06 per µg/m^3 (4.1E-06 per ppb) for lymphoid cancer
- $^{chronic}ESL_{nonthreshold(c)}$ = 4.3 µg/m^3 (2.4 ppb) (ADAF adjusted; rounded to two significant figures)

The long-term ESL for air permit reviews and the evaluation of long-term ambient air monitoring data, set at an excess risk of 1 in 100,000, is the ADAF-adjusted $^{chronic}ESL_{nonthreshold(c)}$ of 4.3 µg/m^3 (2.4 ppb). The ADAF-adjusted URF for lymphoid cancer is 2.3E-06 per µg/m^3 or 4.1E-06 per ppb.

Chapter 4 References

Agency for Toxic Substances and Disease Registry (ATSDR). 1990. Toxicological profile for ethylene oxide.

Ambroise, D, Moulin, JJ, Squinazi, F, Protois, JC, Fontana, JM, Wild, P. 2005. Cancer mortality among municipal pest-control workers. Int Arch Occup Environ Health. 78:387-393.

Austin, S, Sielken, R. 1988. Issues in assessing the carcinogenic hazards of ethylene oxide. Journal of Occupational Medicine. 30(3):236-245.

Bisanti, L, Maggini, M, Raschetti, R, Spila Alegiani, S, Menniti Ippolito, F, Caffari, B, Segnan, N, Ponti, A. 1993. Cancer mortality in ethylene oxide workers. British Journal of Industrial Medicine. 50:317-324.

Bogen, K, Sheehan, P, Valdez-Flores, C, Li, A. 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.

Bratu, AM. 2019. Spectroscopic study of breath ethylene via the mouth and nose. Lasers in Medical Science 34:773-778.

Bruder, C, Bulliard, JL, Germann, S, et al. 2018. Estimating lifetime and 10-year risk of lung cancer. Preventive Medicine Reports 11:125-130.

ChemSpider. 2019. http://www.chemspider.com

Coggon, H, Harris, E, Poole, J, Palmer, K. 2004. Mortality of workers exposed to ethylene oxide extended follow up of a British cohort. Occup Environ Med. 61:358-362.

Crump, KS. 2005. The effect of random error in exposure measurement upon the shape of the exposure response. Dose Response 3(4):456-464.

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

Doll, R, Peto, R, Boreham, J, Suthherland, I. 2005. Mortality from cancer in relation to smoking: 50 years observations on British doctors. British Journal of Cancer 92:426-429.

Donner, EM, Wong, BA, James, RA, Preston, RJ. 2010. Reciprocal translocations in somatic and germ cells of mice chronically exposed by inhalation to ethylene oxide: Implications for risk assessment. Mutagenesis 25:49-55.

"Ethylene Oxide Frequently Asked Questions." American Chemistry Council (ACC). 2018. Retrieved from https://www.americanchemistry.com/EO/Ethylene-Oxide-Frequently-Asked-Questions.pdf

"Ethylene Oxide Sterilization for Medical Devices." United States Food and Drug Administration (USFDA) (April 2019). Retrieved from https://www.fda.gov/medical-devices/general-hospital-devices-and-supplies/ethylene-oxide-sterilization-medical-devices

"Ethylene Oxide. 75-21-8." United States Environmental Protection Agency (USEPA). Retrieved from https://www.epa.gov/sites/production/files/2016-09/documents/ethylene-oxide.pdf

Fenske, J, and Paulson S, 1999. Human breath emissions of VOCs. Journal of the Air & Waste Management Association 49:594-598.

Fondelli, C, Costantini, A, Ercolanelli, M, Pizzo, A, Maltoni, S, Quinn, M. 2007 Exposure to carcinogens and mortality in a cohort of restoration workers of water-damaged library materials following the River Arno flooding in Florence, 4 November 1966. Med Lav. 98, 5:422-431.

Gandini, S, Botteri, E, Iodice, S, et al. 2008. Tobacco smoking and cancer: a meta-analysis. Int J Cancer 122:155-164.

Gardner, M, Coggon, D, Pannett, B, Harris, E. 1989. Workers exposed to ethylene oxide: a follow up study. British Journal of Industrial Medicine. 46:860-865.

Ginevan, ME, & Watkins, DK. 2010. Logarithmic dose transformation in epidemiologic dose-response analysis: use with caution. Regul Toxicol Pharmacol 58(2):336-340.

Greenberg, H, Ott., M, Shore, R. 1990. Men assigned to ethylene oxide production or other ethylene oxide related chemical manufacturing: a mortality study. British Journal of Industrial Medicine. 47:221-230.

Greife, A, Hornung, R, Stayner, G, Steenland, K. 1988. Development of a model for use in estimating exposure to ethylene oxide in a retrospective cohort mortality study. Scand J Work Environ Health. 14(1):29-30.

Hagmar, L, Welinder, H, Lindinl, K, Attewelll, R, Osterman-Golkar, S., Tornqvist, 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.

Hagmar, L, Mikoczy, Z, Welinder, H. 1995. Cancer incidence in Swedish sterilant workers exposed to ethylene oxide. Occupational and Environmental Medicine. 52:154-156.

Health Canada, 2016. Screening Assessment of Ethene (Ethylene), Chemical Abstracts Service Registry Number (CAS RN) 74-85-1, Environment and Climate Change Canada, Health Canada, May 2016.

Hill, AB. 1965. The Environment and Disease: Association or Causation? Proc R Soc Med 58:295-300.

Hogstedt, C, Rohlen, O, Berndtsson, B, Axelson, O, Ehrenberg, L. 1979a. A cohort study of mortality and cancer incidence in ethylene oxide production workers. British Journal of Industrial Medicine. 36:276-280.

Hogstedt, C, Malmqvist, N, Wadman, B. 1979b. Leukemia in workers exposed to ethylene oxide. JAMA. 241(11):1132-1133.

Hogstedt, C., Aringer, L, Gustavsson, A. 1986. Epidemiologic support for ethylene oxide as a cancer-causing agent. JAMA. 255(12):1575-1578.

Hogstedt, C. 1988. Epidemiological studies on ethylene oxide and cancer: an updating. IARC Sci Publ. 89:265-70.

Hornung, R, Greife, A, Stayner, L, Steenland, K, Herrick, R, Elliott, L, Ringenburg, V, Morawetz. 1994. Statistical model for prediction of retrospective exposure to ethylene oxide in an occupational mortality study. American Journal of Industrial Medicine. 25:825-836.

International Agency for Research on Cancer (IARC). 1994. Ethylene: In: IARC, Monographs on the evaluation of carcinogenic risks to humans, Vol 60, Some industrial chemicals, pp-45-71. Lyon, France: International Agency for Research on Cancer.

International Agency for Research on Cancer (IARC). 2008. IARC monographs on the evaluation of carcinogenic risks to humans: 1,3-butadiene, ethylene oxide and vinyl halides (vinyl fluoride, vinyl chloride and vinyl bromide). Lyon, France: International Agency for Research on Cancer.

International Agency for Research on Cancer (IARC). 2012. IARC monographs on the evaluation of carcinogenic risks to humans, No. 100F. Lyon, France: International Agency for Research on Cancer.

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

Jacob, L, Freyn, M, Kalder, M, et al. 2018. Impact of tobacco smoking on the risk of developing 25 different cancers in the UK: a retrospective study of 422,010 patients followed for up to 30 years. Oncotarget 9(25):17420-17429.

Kardos, L, Szeles, G, Gombkoto, G, Szeremi, M, Tompa, A, Adany, R. 2003. Cancer deaths among hospital staff potentially exposed to ethylene oxide an epidemiological analysis. Environmental and Molecular Mutagenesis. 42:59-60.

Kiesselbach, N, Ulm, K, Lange, H-J, Korallus, U. 1990. A multicentre mortality study of workers exposed to ethylene oxide. British Journal of Industrial Medicine. 47:182-188.

Kiran, S, Cocco, P, Mannetje, A, Satta, G, D'Andrea, I, Becker, N, de Sanjose, S, Foretova, L, Staines, A, Kleefeld, S, Maynadie´, M, Nieters, A, Brennan, P, Boffettal, P. 2010. Occupational exposure to ethylene oxide and risk of lymphoma. Epidemiology. 21:905-910.

Kirkeleit, J, Riise, T, Bjørge, T, Christiani, DC. 2013. The healthy worker effect in cancer incidence studies. Am J Epidemiol. 2013; 177(11): 1218-1224.

Kirman, C, Sweeney, L, Teta, M, Sielken, R, Valdez-Flores, C, Albertini, R, Gargas, M. 2004. Addressing nonlinearity in the exposure–response relationship for a genotoxic carcinogen: cancer potency estimates for ethylene oxide. Risk Anal. 24 (5): 1165-1183.

Kirman, C, Hays, S. 2017. Derivation of endogenous equivalent values to support risk assessment and risk management decisions for an endogenous carcinogen. Regulatory Toxicology and Pharmacology. 91:165-172.

Kita, M. 1991. Mortality among workers exposed to ethylene oxide 16-year average follow-up of over 18,000 workers. Journal of Insurance Medicine. 23(3):176-179.

Li, A, Sheehan, P, Valdez-Flores, C, et al. 2019. Updated modeling indicates historical sterilization worker ethylene oxide exposures were higher than assumed in the US EPA IRIS assessment. The Toxicologist 168:1855.

Marsh, GM, Keeton, KA, Riordan, AS, et al. 2019. Ethylene oxide and risk of lympho-hematopoietic cancer and breast cancer: a systematic literature review and meta-analysis. Int Arch Occup Environ Health. 92(7):919-939.

Mikoczy, Z, Tinnerberg, H, Björk, J, Albin, M. 2011. Cancer incidence and mortality in Swedish sterilant workers exposed to ethylene oxide updated cohort study findings 1972-2006. Int. J. Environ. Res. Public Health. 8:2009-2019.

Morgan, R, Claxton, K, Divine, B, Kaplan, S, Harris, V. 1981. Mortality among ethylene oxide workers. Journal of Occupational Medicine. 23(11):767-77.

Mosavi-Jarrahi, A, Mohagheghi, M, Kalaghchi, B, Mousavi-Jarrahi, Y, Kolahi, A, Noori, M. 2009 Estimating the incidence of leukemia attributable to occupational exposure in Iran. Asian Pacific Journal of Cancer Prevention. 10:67-70.

"NATA Frequent Questions. National Air Toxics Assessment." USEPA. Retrieved from https://www.epa.gov/national-air-toxics-assessment/nata-frequent-questions#results

Norman, S, Berlin, J, Soper, K, Middendorf, B, Stolley, P. 1995 Cancer incidence in a group of workers potentially exposed. International Journal of Epidemiology. 24(2):276-284.

Olsen, G, Lacy, S, Bodner, K, Chau, M, Arceneaux, T, Cartmill, J, Ramlow, J, Boswell, J. 1997 Mortality from pancreatic and lymphopoietic cancer among workers in ethylene and propylene chlorohydrin production. Occup Environ Med. 54:592-598.

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

Schulte, PA, Boeniger, M, Walker, JT, Schober, SE, Pereira, MA, Gulati, DK, Wojciechowski, JP, Garza, A, Froelich, R, Strauss, G. 1992. Biologic markers in hospital workers exposed to low levels of ethylene oxide. Mutat Res 278:237-251.

Scientific Committee on Occupational Exposure Limits (SCOEL). 2012. Recommendation from the Scientific Committee on Occupational Exposure Limits for Ethylene Oxide, SCOEL/SUM/160.

Shore, R, Gardner, M, Pannett, B. 1993. Ethylene oxide an assessment of the epidemiological evidence on carcinogenicity. British Journal of Industrial Medicine. 50:971-997.

Sielken Jr, R, Valdes-Flores, C. 2009a. Calculating excess risk with age-dependent adjustment factors and cumulative doses ethylene oxide case study. Regulatory Toxicology and Pharmacology. 55:76-81.

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

Stayner, L, Steenland, K, Greife, A, Hornung, R, Hayes, R, Nowlin, S, Morawetz, J, Ringenburg, V, Elliot, L, Halperin, W. 1993. Exposure response analysis of cancer mortality in a cohort of workers exposed to ethylene oxide. American Journal of Epidemiology. 138:787-798.

Steenland, K, Stayner, L, Greife, A, Halperin, W, Hayes, R, Hornung, R, Nowlin, S. 1991 Mortality among workers exposed to ethylene oxide. The New England Journal of Medicine. 324(20):1402-1407.

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

Steenland, K, Stayner, L, Deddens, J. 2004. Mortality analyses in a cohort of 18 235 ethylene oxide exposed workers: follow up extended from 1987 to 1998. Occup Environ Med. 61:2-7.

Steenland, K, Seals, R, Klein, M, Jinot, J, Kahn, H. 2011. Risk Estimation with Epidemiologic Data When Response Attenuates at High-Exposure Levels. Environmental Health Perspectives. 119(6):831-837.

Swaen, G, Slangen, J, Ott, M, Kusters, E, Van Den Langenbergh, G, Arends, J, Zober, A. 1996. Investigation of a cluster of ten cases of Hodgkin's disease in an occupational setting. Int Arch Occup Environ Health. 68:224-228.

Swaen, G, Burns, C, Teta, J, Bodnar, K, Keenan, D, Bodnar, C. 2009. Mortality study update of ethylene oxide workers in chemical manufacturing: a 15 year update. JOEM. 51(6):714-723.

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

Teta, M, Benson, L, Vitale, J. 1993. Mortality study of ethylene oxide workers in chemical manufacturing a 10 year update. British Journal of Industrial Medicine. 50:704-709.

Teta, J, Sielken Jr, R, Valdez-Flores, C. 1999. Ethylene oxide cancer risk assessment based on epidemiological data application of revised regulatory guidelines. Risk Analysis. 19(6):1135-1155.

TCEQ. 2015. Guidelines to develop toxicity factors. RG-442: Texas Commission on Environmental Quality (TCEQ).

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

United States Environmental Protection Agency (USEPA). 2005b. Supplemental Guidance for Assessing Susceptibility from Early-Life Exposure to Carcinogens. (EPA/630/R-03/003F). Washington, D.C.

United States Environmental Protection Agency (USEPA). 2016. Evaluation of the inhalation carcinogenicity of ethylene oxide (CASRN 75-21-8): In support of summary information on the Integrated Risk Information System (IRIS). EPA/635/R-16/350Fa. Washington, DC, U.S. Environmental Protection Agency, Office of Research and Development, National Center for Environmental Assessment.

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

Valdez-Flores, C, Sielken Jr, R, Teta, J. 2011 Quantitative cancer risk assessment for ethylene oxide inhalation. Arch Toxicol. 85:1189-1193.

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

van Sittert, NJ, Beulink, GD, van Vliet, E, van der Waal, H, 1993. Monitoring occupational exposure to ethylene oxide by the determination of hemoglobin adducts. Environ Health Perspect 99:217-220.

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

Walker, VE, Fennell, TR, Upton, PB, MacNeela, JP, Swenberg, JA. 1993. Molecular dosimetry of DNA and hemoglobin adducts in mice and rats exposed to ethylene oxide. Environ Health Perspect 99:11-17.

Walker, VE, Wu, KY, Upton, PB, Ranasinghe, A, Scheller, N, Cho, MH, Vergnes, JS, Skopek, TR, Swenberg, JA. 2000. Biomarkers of exposure and effect as indicators of potential carcinogenic risk arising from in vivo metabolism of ethylene to ethylene oxide. Carcinogenesis 21:1661-1669.

Wang, TW, Asman, K, Gentzke, AS, et al. 2018. Tobacco Product Use Among Adults — United States, 2017. MMWR Morb Mortal Wkly Rep 67:1225-1232.

World Health Organization (WHO). 2003. Concise international chemical assessment document: Ethylene oxide. In Concise International Chemical Assessment: Ethylene Oxide. Geneva, Switzerland: World Health Organization, International Programme on Chemical Safety.

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

Appendix 1 Systematic Review and Evidence Integration

A1.1 Problem Formulation and Protocol

Problem formulation identifies and defines the causal questions and describes the extent of the evaluation. These questions structured the systematic review for EtO:

- What are the physical and chemical properties of EtO?
- What is the critical effect following exposure to EtO?
- Are there sensitive subpopulations?
- What is the mode of action (MOA)?
- Does route of exposure play a role?
- Is EtO carcinogenic, and if so, is it carcinogenic by a specific route of exposure?

Protocol development is another important aspect in the initial process. A protocol is typically developed around a PECO statement: Populations, Exposure, Comparator/Control, and Outcomes. These identifiers are used to lay out the framework for the literature search and inclusion/exclusion criteria. The PECO statement for EtO followed these criteria:

Table 17: PECO Statement used by the TCEQ to Develop Toxicity Factors for EtO

Population	General human population and any relevant sensitive subpopulations, animals, and vegetation
Exposure	Exposure to EtO, surrogates with demonstrated similar MOAs, and any identified metabolites
Comparator/ Control	Populations exposed to concentrations below the concentration that causes the most sensitive critical effect
Outcome(s)	The most sensitive critical effect directly related to EtO exposure

The protocol used for the systematic review and the development of toxicity factors for EtO is as follows:

1. Identify the chemical of interest and define the causal questions
2. Conduct a systematic review
    a. Conduct a systematic literature search
    b. Identify the inclusion/exclusion criteria
    c. Extract the relevant data from each data stream (human, animal, mechanistic)
    d. Assess the study quality and conduct a risk of bias analysis
    e. Weigh the evidence in each data stream and then integrate the evidence across the data streams

   f. Rate the confidence in the evidence
3. Derive toxicity factors (TCEQ 2015)
   a. Review the essential data, including chemical/physical properties and selected key studies from the systematic review
   b. Conduct MOA analysis
   c. Choose the appropriate dose metric considering toxicokinetics and MOA
   d. Select critical effect, based on human equivalent exposure considering each key study
   e. Extrapolate from the adjusted POD to lower exposures based on MOA analysis

A1.2 Systematic Literature Review and Study Selection

As a first step, publically available databases were searched using explicitly stated search criteria. Please see TCEQ (2015) for a list of available databases that were searched. The search terms used in literature review for EtO, along with the number of results from PubMed, are found in Table 18. Additional references were also identified using the reference sections from some of the selected studies. This literature review was conducted in December 2018, and therefore studies published after this date were not available at the time of the review.

Table 18: Search Strings used in the Literature Review of EtO

Search Term/String	PubMed Results
ethylene oxide	9,626
“ethylene oxide”	7,478
“ethylene oxide” OR oxirane	10,374
“ethylene oxide” OR oxirane OR 75-21-8	10,374

These 10,374 studies were imported into the desktop application SWIFT-Review by Sciome and briefly searched to ensure that the key studies used in several other reviews were present in the data set. The data set was further narrowed down using the tag levels created by the SWIFT-Review software. The tags used and the number of studies that each tag removed can be found in Table 19.

Table 19: SWIFT-Review Tags and Results

Data Set/Tag	Number of Studies
Initial PubMed Search	10,374
Tag – Health Outcomes, any (excluded studies with no tag)	7,468
Tag – Evidence Stream, any (excluded studies with no tag)	4,914
Tag – MeSH Chemicals, only Ethylene Oxide (excluded everything else)	1,520

Additionally, several governmental and private sector organizations were searched for published literature and toxicity values for EtO (Table 20), and the available documents along with their relevant references were added to the pool of selected material as needed.

Table 20: Available Reviews and Inhalation Toxicity Values for EtO

Organization	Year	Toxicity Value
Agency for Toxic Substances and Disease Registry (ATSDR) Toxicological Profiles	1990	Intermediate MRL*
Integrated Risk Information System (IRIS) USEPA	2016	Inhalation Unit Risk
Office of Environmental Health Hazard Assessment (OEHHA) CalEPA	2000	Chronic REL* Inhalation Slope Factor

MRL – minimal risk level, REL – reference exposure level

Following this initial review, specific inclusion and exclusion criteria were used to narrow down the pool of available data. The criteria along with examples of the kinds of studies that were excluded can be found in Table 21.

Table 21: Inclusion/Exclusion Criteria used in the Review of EtO

Study Type	Inclusion Criteria	Exclusion Criteria
General	Complete study available for review	- Only abstract is available - Study in a language other than English - Unpublished report/unable to retrieve
	Study contains original data or utilizes existing data in a novel way	- Study is a review article or meta-analysis - Study comments on a previous method without providing a sufficient alternative
	Exposure concentration is known or can be reasonably estimated	- Exposure concentration unknown - Exposure environment/conditions unsuitable to concentration estimation
	Study examines effects related to chemical exposure	- Study measures concentration in air, factories, etc. - Study does not examine health effects
	Study focused on the chemical of concern	- Study examined mixture effects - Study on treatment following EtO exposure
	Route of exposure is relevant to exposure and toxicity factor development	- Exposure through i.v., i.p., or subcutaneous injection - Study examining oral or dermal exposure
Animal	Relevant animal model and endpoints examined	- Study used non-mammalian animal models - Endpoint studied not relevant to human health - Endpoint not applicable to toxicity factor development
	Appropriate study populations and methods were used	- Study lacked appropriate numbers or doses - Exposure method unsuitable for dose-response
Human/Epi	Relevant endpoints examined	- Study focused solely on cytotoxicity - Study only measured sister chromatid exchanges (SECs), protein adducts, or chromosomal changes
	Study populations allowed for significant findings and follow ups	- Case studies examining single high-dose exposures - Studies without appropriate follow-up studies - Historical studies that have been updated

i.v. – intravenous, i.p. – intraperitoneal

Studies were then divided into groups by evidence stream (i.e. human, animal) and effect group (i.e., acute, chronic non-carcinogenic, carcinogenic). For the purposes of this DSD, only the human carcinogenic/epidemiologic data were considered for several reasons:

1. In order to expedite the process, it was decided that only a health-based chronic carcinogenic toxicity factor would be derived for EtO in this DSD. Other toxicity factors (i.e. health- and welfare-based acute and chronic non-carcinogenic) may be evaluated at a later date with an additional systematic review picking up where this systematic review left off.
2. Sufficient human data exist for EtO such that animal data, although used to strengthen the carcinogenicity class, would not be used to derive a chronic carcinogenic toxicity factor. TCEQ (2015) states that in general, human data are preferred over animal data when developing toxicity factors.
3. Similarly, mechanistic data remain supportive (e.g., MOA), but not useful as a basis in the derivation of a chronic carcinogenic toxicity factor.
4. And finally, human data looking solely at cytotoxicity, sister chromatid exchanges, or chromosomal abnormalities were considered useful in developing the MOA of EtO, but not useful as a basis for derivation of a health-based toxicity factor.

After full text review and screening with the inclusion/exclusion criteria listed above, eight human carcinogenic studies were identified for further use in this systematic review. Several human studies (directly or indirectly related to carcinogenicity) were reviewed and later excluded due to various reasons (Table 22).

Table 22: Excluded Human Studies Related to Carcinogenicity

Reason for Exclusion	Study	
No exposure or dose-response information available to directly derive a toxicity factor (Not useful in the development of a carcinogenic-based toxicity factor)	Ambroise et al., 2005 Austin and Sielken, 1988 Bisanti et al., 1993 Coggon et al., 2004 Fondelli et al., 2007 Gardner et al., 1989 Greenburg et al., 1990 Greife et al., 1988 Hagmar et al., 1991 Kardos et al., 2003	Kiesselbach et al., 1990 Kiran et al., 2010 Kirman and Hays, 2017 Morgan et al., 1981 Mosavi-Jarrahi et al., 2009 Norman et al., 1995 Olsen et al., 1997 Swaen et al., 1996 Wong and Trent, 1993
Follow up study available	Greenberg et al., 1990 Hagmar et al., 1995 Hogstedt et al., 1979a Hogstedt et al., 1986	Stayner et al., 1993 Steenland et al., 1991 Teta et al., 1993
Review, methods, or case study	Hogstedt et al., 1979b Hornung et al., 1994 Kita, 1991 Shore et al., 1993 Sielken and Valdez-Flores, 2009a	Sielken and Valdez-Flores, 2009b Steenland et al., 2011 Valdez-Flores et al., 2011 Valdez-Flores and Sielken, 2013

A1.3 Data Extraction

Each of the identified studies was reviewed in detail and the primary data were extracted for potential use in the development of the chronic carcinogenic toxicity factor in this DSD (Table 23).

Table 23: Data Extraction from Epidemiological Studies

Study (cohort)	Size	Exposure Measurement	Tumor Type(s)	Notable Results[1]	Notes
Hogstedt 1988 (Swedish, chemical)	539 m 170 f	Years of employment, 1-9 years, ≥ 10 years	Stomach	SMRs – 597, 608	Exposure estimates conducted in original study but not presented here.
			Blood/Lymphatic	SMRs – 380, 330	
			Leukemia	SMRs – 322, 880	
Kirman 2004 (NIOSH + UCC)	18,254 (NIOSH) (55% m, 45% f) 1,896 m (UCC)	ppm-years, 7.4, 64.8, 187.4, 477.7	Leukemia	POD-ED_{001} estimated at 265 ppm-years, URFs: linear 4.5×10^{-7} /μg/m^3 Quadratic 4.5×10^{-8} /μg/m^3 (no lag or latency periods)	Concentration at 1×10^{-5} cancer risk: Linear – 22 μg/m^3 (12 ppb) Quadratic – 222 μg/m^3 (120 ppb) Nonlinear – 37 μg/m^3 (21 ppb)
Mikoczy 2011 (Swedish, sterilant)	862 m 1,309 f	ppm-years, 0-0.13, 0.14-0.21, ≥ 0.22	Breast	SIRs – 0.52, 1.06, 1.12	Compared with/out 15-year latency and between follow-ups
			LHN	SIRs – 1.35, 1.32, 1.08	
Steenland 2003 (NIOSH)	7,576 f (5,139 f interviewed)	ppm-days, 0, >0-647, 647-2026, 2026-4919, 4919-14620, 14620+	Breast (Compared to US population)	SIRs – 0.88, 0.77, 0.77, 0.94, 0.83, 1.27 (15-year lag, cumulative)	Subset of the NIOSH cohort, multiple other comparisons presented, including cumulative, categorical, and log cumulative exposure, positive trends for continuous exposure, duration of exposure, and log of cumulative exposure. Overall SMR for NIOSH cohort for breast cancer is 0.99. Exposure-response analysis showed highest group SMR of 1.27, with 20-year lag increased to 2.07 (95% CI: 1.0-3.54)
			Breast (Compared to study population, whole cohort)	Odds Ratios – 1.00, 1.07, 1.00, 1.24, 1.17, 1.74* (15-year lag, categorical, cumulative)	
			Breast (Compared to study population, only interviewed cohort)	Odds Ratios – 1.00, 1.06, 0.99, 1.24, 1.42, 1.87* (15-year lag, categorical, cumulative)	

Study (cohort)	Size	Exposure Measurement	Tumor Type(s)	Notable Results[1]	Notes
Steenland 2004 (NIOSH)	7,645 m 9,885 f	ppm-days, 0, >0-1199, 1200-3679, 3680-13499, 13500+	NHL	SMRs – 2.09, 0.61, 0.88, 0.79, 2.37* m, 10-year lag, cumulative	Multiple other comparisons presented, including cumulative, categorical, and log cumulative exposure, 10, 15, and 20-year lag, positive trend for lymphoid tumors
		ppm-days, 0, >0-646, 647-2779, 2780-12321, 12322+	Breast	SMRs –0.80, 1.05, 1.01, 1.15, 2.07* f, 20-year lag, cumulative	
Swaen 2009 (UCC)	2,063 m	ppm-years, 0-15, 15-65, 65+	None	Authors state no long-term carcinogenic effects associated with EtO exposure	Cohort experienced more than twice the average estimated cumulative exposure compared to NIOSH cohort
Teta 1999 (multiple reviewed, dose-response done for NIOSH and UCC)	Multiple, meta-analysis 8,214 m & 10,040 f (NIOSH) 1,896 m (UCC)	ppm-years, 0, 0-33, 33-125, 125-285, >285	Lymphoid (lymphocytic leukemia and NHL)	Added Risk (environmental) UCC – none NIOSH – 10^{-8} – 10^{-5} /ppb	Compared 0 and 10-year latency, and 0 and 5y lag periods, POD-ED_{001} values ranged from 0.81-1.58 ppm assuming a 10-year latency and a 5-year lag period. POD-ED_{001} of 0.81 ppm gives a URF of 0.12/ppm, and a concentration at 1×10^{-5} cancer risk of 0.083 ppb (0.15 μg/m^3)
			Leukemia	Added Risk (environmental) UCC – 10^{-12} – 10^{-6} /ppb NIOSH – 10^{-15} – 10^{-6} /ppb	
Valdez-Flores 2010 (NIOSH + UCC)	7,634 m & 9,859 f (NIOSH) 2,063 m (UCC)	ppm-days, dose ranges varied by endpoint	Examined 12 cancer endpoints in 6 subcohorts	No statistically significant increases in SMRs, trends, cumulative continuous, or categorical exposure.	No heterogeneity between dose-response models of the two major cohorts and the pooled study, combining increases the power.

[1] Due to space constraints, only notable results are presented here. See individual studies for a more in-depth review.

* Denotes significance, confidence interval does not include 1

SMR – Standardized mortality ratio, SIR – Standardized Incidence Ratio, NHL – Non-Hodgkin's Lymphoma, LHN – Lymphohematopoietic Neoplasms, m – males, f – females

A1.4 Study Quality and Risk of Bias (ROB)

Each of the selected studies was evaluated for study quality and ROB based on a number of attributes determined prior to this review. For this review, study quality methods were adapted from the USEPA version of the Health Assessment Workspace Collaboration (HAWC) online software. For epidemiology studies, seven evaluation domains are used to critically assess different aspects of study design and conduct relating to reporting, risk of bias and study sensitivity. Each domain receives a score of Good, Adequate, Deficient, Critically Deficient, or Not Reported, and once all domains are evaluated, a confidence rating of High, Medium, or Low confidence or Uninformative is assigned to each study. The evaluated domains and explanations can be found in Table 24, while the general guidance for scoring each of the studies can be found in Table 25 and Table 26.

Table 24: Study Quality Domains for Epidemiology Studies (taken from HAWC)

Domain	Study Design Questions and Aspects
Selection and Performance/ Participant Selection	**Is there evidence that selection into or out of the study (or analysis sample) was jointly related to exposure and to outcome?** Study design, where and when was the study conducted, and who was included? Recruitment process, exclusion and inclusion criteria, type of controls, total eligible, comparison between participants and nonparticipants (or followed and not followed), final analysis group. Does the study include potential vulnerable/susceptible groups or life stages?
Exposure Methods/ Measures	**Does the exposure measure reliably distinguish between levels of exposure in a time window considered most relevant for a causal effect with respect to the development of the outcome?** Source(s) of exposure (consumer products, occupational, an industrial accident) and source(s) of exposure data, blinding to outcome, level of detail for job history data, when measurements were taken, type of biomarker(s), assay information, reliability data from repeat measures studies, validation studies.
Outcome Methods/Results Presentation	**Does the outcome measure reliably distinguish the presence or absence (or degree of severity) of the outcome?** Source of outcome (effect) measure, blinding to exposure status or level, how measured/classified, incident versus prevalent disease, evidence from validation studies, prevalence (or distribution summary statistics for continuous measures).
Confounding	**Is confounding of the effect of the exposure unlikely?** Background research on key confounders for specific populations or settings; participant characteristic data, by group; strategy/approach for consideration of potential confounding; strength of associations between exposure and potential confounders and between potential confounders and outcome; degree of exposure to the confounder in the population.

Domain	Study Design Questions and Aspects
Analysis	**Does the analysis strategy and presentation convey the necessary familiarity with the data and assumptions?** Extent (and if applicable, treatment) of missing data for exposure, outcome, and confounders, approach to modeling, classification of exposure and outcome variables (continuous versus categorical), testing of assumptions, sample size for specific analyses, relevant sensitivity analyses.
Selective Reporting	**Is there concern for selective reporting?** Are results presented with adequate detail for all the endpoints of interest? Are results presented for the full sample as well as for specified subgroups? Were stratified analyses (effect modification) motivated by a specific hypothesis?
Sensitivity	**Are there concerns for study sensitivity?** What exposure range is spanned in this study? What are the ages of participants (e.g., not too young in studies of pubertal development)? What is the length of follow-up (for outcomes with long latency periods)? Choice of referent group and the level of exposure contrast between groups (i.e., the extent to which the 'unexposed group' is truly unexposed, and the prevalence of exposure in the group designated as 'exposed'). Is the study relevant to the exposure and outcome of interest?
Overall Study Confidence	**Once the evaluation domains have been classified, these ratings will be combined to reach an overall study confidence classification of High, Medium, Low, or Uninformative.** This classification will be based on the classifications in the evaluation domains and will include consideration of the likely impact of the noted deficiencies in bias and sensitivity on the results.

Table 25: Study Quality Domain Scoring

Score	Reasoning
+ +	**Good** – Study meets or exceeds domain properties, may have minor deficiencies but none that would affect the outcome of the study or the development of toxicity factors.
+	**Adequate** – Study meets most of the domain properties, may have some deficiencies but none are severe nor are expected to have a serious effect on the development of toxicity factors.
-	**Deficient** – Study has one or more deficiencies that are likely to affect the outcome of the study or the development of toxicity factors, but development may still occur with some added uncertainty.
- -	**Critically Deficient** – Study has serious deficiencies that would severely inhibit the development of toxicity factors. These studies are typically classified as "uninformative" unless a detailed explanation otherwise is provided.
NR	**Not Reported** – Domain properties are not provided in the study or referred to in previous author's studies. Depending on the domain and type of study, these studies should be carefully considered prior to use.

Table 26: Study Quality Confidence Rating Scoring

Score	Reasoning
+ +	**High** – Overall a well conducted study, no serious deficiencies identified, no concern for issues with sensitivity or ROB, most domains should be scored good or adequate.
+	**Medium** – Some deficiencies may be noted, but nothing that would cause significant concern for issues with sensitivity or ROB, most domains should be scored adequate.
-	**Low** – Deficiencies noted, some severe, and some concern over bias or sensitivity that may impact the assessment, study has domains that scored deficient.
- -	**Uninformative** – Severe deficiencies that would seriously impact the assessment, study is typically unusable for toxicity factor development without a detailed explanation. Any study with a domain listed as "Critically Deficient" should be considered for this category.

Scoring for each of the included studies can be found in Table 27. Each reviewer scored the included studies independently, then came together to agree on a single score for each domain/study (individual scoring not shown).

Table 27: Study Quality and ROB Scoring Visual

Domain/Study	Hogstedt 1988	Kirman 2004	Mikoczy 2011	Steenland 2003	Steenland 2004	Swaen 2009	Teta 1999	Valdez-Flores 2010
Selection and Performance/Participant Selection	+	+ +	+	+	+ +	+	+ +	+ +
Exposure Methods/Measures	-	+	-	+	+	-	+	+
Outcome Methods/Results Presentation	+	+	+ +	+	+	+	+	+ +
Confounding	-	+	-	+ +	+	+	+	+
Analysis	+	+	+	+	+ +	+	+	+ +
Selective Reporting	+	+	+	+	+	+	+	+
Sensitivity	-	+	-	+	+	+	+	+
Overall Study Confidence	-	+	+	+	+	+	+	+

A1.5 Evidence Integration

After addressing the study quality and ROB for each of the selected studies, the primary information from each of the studies was compiled together and each study was assessed for use as a key, supporting, or informative study (Table 28).

Table 28: Evidence Integration Table for Human Studies

Study	Cohort	Type	Reasoning
Hogstedt 1988	Swedish chemical workers	Informative	- Relatively small cohort with little information on co-exposures - Exposure concentrations or estimations not provided - Primary cohort to show increased leukemia mortality rates - Also presented increased stomach and blood/lymphatic cancer
Kirman 2004	NIOSH + UCC	Supporting	- Combined data from two largest cohorts and examined leukemia and lymphoid tumor mortality data - Provided results for several different extrapolation methods - Selected a single outcome and POD to carry through
Mikoczy 2011	Swedish sterilant workers	Informative	- Relatively small cohort with little exposure information presented - Healthy worker effect likely influenced the results - Non-significant increases in leukemia, NHL, and lymphohematopoietic cancer mortality - Significant increases in the rate ratios of breast cancer in the two highest exposure groups
Steenland 2003	NIOSH (females only)	Informative	- Subset of the largest cohort study available, additional nested case-control using subjects who answered personal interviews - Examined breast cancer mortality and incidence data - Positive trend for increased incidence, but not significantly increased
Steenland 2004	NIOSH	Supporting	- Update to the largest EtO-exposed cohort data available - Focused mainly on hematopoietic and breast cancers, and examined various exposure variables and lag periods - No significantly increased cancer incidences, but a positive trend observed for lymphoid tumors (males, 15-year lag)

Study	Cohort	Type	Reasoning
Swaen 2009	UCC	Supporting	- Although a relatively smaller cohort, the strength of the update was made up for in the length of follow-up and number of deaths - Little to no exposure monitoring data available, estimates made from work history - Examined a wide array of cancer types but no lag/latency periods - No cancer associations observed
Teta 1999	Meta-analysis, NIOSH, UCC	Supporting	- Very basic meta-analysis of 10 EtO cohorts but lacked dose-response data, detailed analysis on individual NIOSH and UCC cohorts only - Examined lymphoid and leukemia rates with various lags and latency periods and control groups using Poisson regression - UCC cohort showed no added risk, while NIOSH cohort predictions were in the range of 10^{-7} to 10^{-5} at 1 ppb environmental exposures
Valdez-Flores 2010	NIOSH + UCC	Key	- Combined most recent data from the UCC and NIOSH cohorts - Examined 12 cancer endpoints (breast, leukemia, lymphoid, etc.) and 6 sub-cohorts (NIOSH males, females, UCC males, etc.) using Cox proportional analyses without latency/lag periods - No statistically significantly increasing SMRs or trends in any of the cancer endpoints examined

After final review of the included studies, the Valdez-Flores et al. (2010) study had the most thorough and complete analysis (e.g., data from both the NIOSH and UCC cohorts, multiple cancer endpoints examined) and was therefore selected as the key study. While the Valdez-Flores et al. (2010) study also utilized a default lifetime duration (70 years) consistent with TCEQ guidance (TCEQ 2015), there were aspects that were not ideal, such as the lack of exposure lags. So rather than select a POD from the key study, the Toxicology, Risk Assessment, and Research Division (TRARD) selected data from both cohorts evaluated in the study (i.e., the NIOSH and UCC cohorts) as the key epidemiological data and conducted an independent assessment using the same approach but with supplemental analyses (e.g., the evaluation of various exposure lags). Selection of data from the NIOSH and UCC cohorts as the key epidemiological data and use of specific, TCEQ-directed dose-response assessment analyses (rather than selection of a study POD) provide the best basis for a carcinogenic assessment of EtO for several reasons:

1. Both the NIOSH and UCC cohorts have adequate size, exposure information, and follow-up, making consideration of all the data ideal for toxicity factor development (e.g., weight of evidence, more analyses to consider).
2. The Valdez-Flores et al. (2010) study makes use of the Cox Proportional Hazard model, a standard model that the TRARD has used previously in dose-response assessments (also considered by USEPA 2016).
3. Although Valdez-Flores et al. (2010) did not include exposure lag results in their publication, supplemental analyses involving a reassessment of the data using various exposure lags allow for the consideration of even more assessment results in the DSD.
4. Additionally, since published in 2010, an update to the UCC data through 2013 has become available to the first author of the Valdez-Flores et al. (2010) study (submitted for publication), who the TCEQ contracted with to perform supplemental analyses; consequently, results from the new study update with a longer follow-up period can also be included in the DSD.
5. Unlike USEPA (2016) that uses a lifetime exposure duration value of 85 years, the TCEQ-directed dose-response analyses use a standard default of 70 years consistent with TCEQ guidance (TCEQ 2015).
6. And finally, conducting these new analyses will allow for the appropriate consideration of model fit to the individual data (rather than the categorical data) for the model assessment ultimately selected by the TRARD.

A1.6 Confidence Rating

Table 29 provides scoring criteria to rate the confidence and uncertainty for each aspect or element of the toxicity assessment. The table provides the name of the element and the magnitude of the confidence in each element using a qualitative ranking system of low, medium, or high confidence. Table 30 displays the overall confidence in the EtO carcinogenic assessment. Once the noncarcinogenic assessments are completed for EtO, the confidence rating will be updated to cover the entire assessment.

Table 29: Confidence Scoring Criteria for EtO Carcinogenic Assessment

Element	Low	Medium	High
Database Completeness	Only a single study or a few low-quality studies were available.	Several studies were available, but some important studies were missing.	Several high-quality studies were available to select from.
Systematic Review	A systematic approach was not used.	A systematic approach was considered and some methods were applied, but a full review was not conducted	A systematic approach was used in study evaluation and clear criteria were established for judgment
Key Study Quality	Selected study has deficiencies, but was still considered useful	Selected study was reasonably well done but some restrictions must be considered	Selected study was well done and can be used without restriction
Critical effect	Critical effect or dose-response curve was moderate to severe. MOA information was not available.	Critical effect was moderate; other studies were deemed necessary to determine the critical effect.	Critical effect was minimal, or the confidence in the critical effect was high. MOA information was available.
Relevance of Critical Effect	Critical effect was only presumed to be relevant for the general population; MOA was not known for the critical effect.	Critical effect appeared to be relevant for the general population. MOA was known for the critical effect and possibly relevant to humans.	Critical effect based on a human study or matches observed human experience; MOA was well understood so critical effect was assumed relevant.
Point of Departure (POD)	Many uncertainties exist in POD; only a few dose groups; no dose-response modeling was used.	Some uncertainty exists in POD; few dose groups; difference between confidence limits was large.	Basis for POD well understood; multiple dose groups, dose-response modeling was conducted.
Sensitive Populations	Many uncertainties on sensitive population(s) existed and were not addressed.	Information on sensitive population(s) was not known but default procedures are presumed to be conservative.	Human data on sensitive populations were available and uncertainties were addressed.
Peer Review	Limited or no peer review; disregarded comments would significantly change risk value; no independent check	Adequate peer review. Most substantive comments addressed; disregarded comments would not significantly change value	High quality panel peer review with appropriate experts; all substantive comments addressed as per independent check
Toxicity Value Comparison	Relevant risk values show a greater than 10-fold difference without justification.	Some relevant risk values agreed within 3-fold of each other, others disagreed within 10-fold without justification.	All relevant risk values agreed within 3-fold of each other or there was sufficient justification for differences.

Table 30: Confidence in the Toxicity Assessment

Element	Score	Basis
Database Completeness	Medium	- Several occupational cohorts (i.e., preferred human data) and animal studies available - Evidence of carcinogenic effects found in both human epidemiological and animal studies - However, estimated exposures are based on incomplete information, are remarkably high, and are not in/near lower range of interest (i.e., not environmentally relevant)
Systematic Review	High	- Systematic review conducted
Key Study Quality	High	- Well-conducted study of two cohorts and multiple cancer endpoints with standard Cox proportional hazards modeling but lacked the use of a lag period - Reassessment of these key epidemiological data utilizing multiple exposure lags and new UCC cohort data allowed for informative supplemental and updated analyses
Critical effect	Low	- Human data not conclusive despite remarkably high exposure (e.g., results vary between studies) - Model (slope > 0) not statistically significantly different than the null model (slope = 0) at the 5% significance level
Relevance of Critical Effect	Medium	- Assumed relevant although general population exposed to levels orders of magnitude lower than the occupational study wherein lymphoid cancer was statistically increased only in the highest cumulative exposure group
Point of Departure (POD)	High	- Cox Proportional Hazard model used - Modeling results demonstrated to be predictive
Sensitive Populations	Medium	- No specific data on sensitive subpopulations - Default ADAFs were applied to account for potentially increased susceptibility in children due to early-life exposure
Peer Review	Medium	- DSD proposed for public comment and reviewed by a consulting academic statistician and subject matter expert in regard to potential statistical issues at TCEQ's direction
Toxicity Value Comparison	High	- TCEQ Chronic ESL based on lymphoid cancer mortality is 4,000 times higher than the USEPA value based on lymphoid/breast cancer incidence at the same excess risk level (1E-05)

Element	Score	Basis
		- TCEQ's approach is supported by multiple lines of evidence as discussed in the DSD, whereas USEPA's non-standard approach is not - Extensive comparisons, calculations, and explanations as to the differences and errors in USEPA's methods are included in the DSD (e.g., USEPA's model assessment is demonstrated to be statistically significantly over-predictive)

Confidence Scoring Summary			
Not Evaluated	**Low Confidence**	**Medium Confidence**	**High Confidence**
	Critical Effect	Database Completeness	Systematic Review
		Relevance of Critical Effect	Key Study Quality
		Sensitive Populations	Point of Departure
		Peer Review	Toxicity Value Comparison

Appendix 2 Reality Check of Epidemiological Exposure-Response Model Results for EtO and Lymphoid Cancer Mortality

USEPA fit several alternative parametric models for lymphoid cancer mortality in the NIOSH cohort and compared the predicted rate ratios by each model with non-parametric estimates of rate ratios. USEPA used the visual comparison of the parametric and non-parametric rate ratios as one of their criteria to select their parametric model. A more robust comparison is to see how reasonable the parametric models are when comparing what the models predict in terms of lymphoid cancer deaths versus the actual number of deaths in the NIOSH cohort. A good (i.e., reasonably accurate) parametric model should predict the observed number of lymphoid cancer deaths with some confidence (e.g., the observed number of lymphoid cancer deaths in the NIOSH cohort should be inside a 95% confidence interval of the estimated number of lymphoid cancer deaths).

Here, some of the USEPA models and one model developed by Sielken & Associates (S&A) were used to check whether the models were reasonable; that is, whether the models predicted within a margin of error, the number of lymphoid cancer deaths in the NIOSH cohort. The estimated number of lymphoid cancer deaths for a specific model for the rate ratios were calculated using age-, sex-, race-, and calendar-year specific background hazard rates. Sections C and D of this appendix illustrates how the calculations to determine the number of expected deaths for each model were performed with methodology used in the calculation of standard mortality ratios (SMRs). The SMR is a measure that shows the ratio of observed to expected number of deaths in the cohort. Similarly, the 100(1-α)% confidence interval on the SMR is a confidence interval on the ratio of observed to expected number of deaths in the cohort.

Herein, the inverse of the SMR is used as a measure of over-prediction or under-prediction of the actual number of observed deaths. That is, the inverse of the SMR (SMR^{-1}) is the ratio of expected to observed number of deaths. Similarly, the inverse of the confidence limits of the 100(1-α)% confidence interval on the SMR result in a 100(1-α)% confidence interval on the inverse of the SMR. In turn, using the SMR^{-1} and its 100(1-α)% confidence interval, a 100(1-α)% confidence interval on the expected or predicted number of deaths can be easily calculated. Using this confidence interval on the predicted number of deaths can then be compared with the observed number of deaths. If the observed number of deaths is inside the 100(1-α)% confidence interval, then the expected number and observed number of deaths are not statistically significantly different at the α% significance level. If the observed number of deaths is below the lower end or above the upper end of the 100(1-α)% confidence interval, then the expected number is statistically significantly different than the observed number of deaths at the α% significance level.

At issue is the predictiveness (or lack thereof) of the model assessments ultimately used by USEPA and the TCEQ. *There is no fairer evaluation of the predictiveness of a given model assessment than direct numerical comparisons of the specific model's predictions to the reality of the dose-response data.* **Upon performing this evaluation, the sections below show that only the log-linear model (standard Cox proportional hazards model; TCEQ's selected model) and the best estimates of the linear model predict the number of observed lymphoid deaths in the NIOSH cohort with 95% confidence.** *By contrast, the model chosen by USEPA (i.e., the linear two-piece spline model with the "knot" at 1,600 ppm-days; 15-year exposure lag) statistically significantly over-estimates (statistically significant at the 5% significance level) the number of observed lymphoid cancer deaths* (even after restricting those models to assume zero increase in the rate ratio for cumulative exposures above the knot).

A2.1 Predicted Versus Observed Number of Lymphoid Cancer Deaths in the NIOSH Cohort

Table 31 and Figure 13 below shows the predicted number of lymphoid cancer deaths in the NIOSH cohort for male and female NIOSH workers using several different EtO exposure-response models. There are 53 lymphoid cancer deaths in the NIOSH cohort (brown horizontal line in Figure 13). Several exposure-response models fit to the NIOSH data were used to estimate the number of lymphoid cancer deaths that the model would predict in the NIOSH cohort, if the fitted model were true. The maximum likelihood estimates of the model as well as the upper 95% confidence limit on the model parameters were used to obtain the predicted number of deaths. In addition to calculating the expected number of deaths predicted by each model and its upper bound on the slope, a 95% confidence interval in the predicted number of deaths was derived using a confidence interval for the ratio of the predicted to the observed number of lymphoid cancer deaths in the NIOSH cohort.

The 95% confidence intervals for the number of lymphoid cancer deaths predicted by the log-linear models and its upper bounds (Cox proportional hazards model, models 1, 2, 3, and 4) include the number of lymphoid cancer deaths actually observed (53) in the NIOSH cohort. The 95% confidence interval for the number of lymphoid cancer deaths predicted by the best estimate of the linear model (model 5) also includes the number of lymphoid cancer deaths actually observed in the NIOSH cohort, but the upper bound of the linear model (model 6) statistically significantly over-predicts the observed number of lymphoid cancer deaths.

Models 7, 8, 9 and 10 are USEPA's two-piece spline models. Every two-piece spline model estimate of the lymphoid cancer deaths in the NIOSH cohort statistically significantly over-predicts the actual number of lymphoid cancer deaths in the NIOSH cohort. For comparison purposes, Models 11, 12, 13 and 14 are USEPA two-piece spline models restrained by setting the slope after the knot equal to zero (i.e., the rate ratio increases with cumulative exposure up to the knot and stays flat after the knot). Even for these restrained two-piece spline

models, for both the MLE and 95% UCL, every model estimate of the lymphoid deaths in the NIOSH cohort statistically significantly over-predicts the actual number of lymphoid deaths in the NIOSH cohort.

Table 31: USEPA's Selected Model Assessment Statistically Significantly Over-Predicts Lymphoid Cancer Mortalities

Model	Slope Parameter (per ppm-day)	Predicted if the Model were True	100% x Ratio: Predicted / Observed	95% CI on Predicted if the Model were True
Background (No Model)	n/a	50.39	95.1%	(38.5, 67.3)
1. S&A – Loglinear – 15-yr lag (MLE)[1]	2.81E-06	52.42	98.9%	(40.1, 70.0)
2. S&A – Loglinear – 15-yr lag (95% UCL)[1] **- TCEQ Adopted**	7.17E-06	58.75	110.8%	(44.9, 78.4)
3. USEPA - Loglinear - 15-yr Lag (MLE)[1] USEPA Table 4-2	4.74E-06[2]	54.52	102.9%	(41.7, 72.8)
4. USEPA - Loglinear - 15-yr Lag (95% UCL)[1] USEPA Table 4-2	1.03E-05[3]	66.41	125.3%	(50.8, 88.7)
5. USEPA - Linear - 15-yr Lag (MLE) USEPA Table D-36	1.23E-05[4]	57.58	108.6%	(44.0, 76.9)
6. USEPA - Linear - 15-yr Lag (95% UCL) USEPA Table D-36	**4.71E-05**[5]	**77.3**	**145.8%**	**(59.1, 103.2)**
USEPA Spline Model with Knot at 1,600 ppm-days				
7. USEPA – Loglinear Spline – 15-yr lag (MLE) – USEPA Table 4-4 Knot @ 1,600 ppm-days	**4.89E-04**[6]	**88.24**	**166.5%**	**(67.5, 117.8)**
8. USEPA – Loglinear Spline – 15-yr lag (95% UCL) – USEPA Table 4-4 Knot @ 1,600 ppm-days	**9.08E-04**[7]	**144.15**	**272.0%**	**(110.2, 192.5)**
9. USEPA – Linear Spline – 15-yr lag (MLE) – USEPA Table 4-4 Knot @ 1,600 ppm-days	**7.58E-04**[8]	**91.69**	**173.0%**	**(70.1, 122.4)**
10. USEPA – Linear Spline – 15-yr lag (95% UCL) – USEPA Table 4-4 Knot @ 1,600 ppm-days - USEPA Selected	**1.80E-03**[9]	**141.09**	**266.2%**	**(107.9, 188.4)**
Results using above USEPA models but assuming that slope for RR is zero after the "knot"				

Model	Slope Parameter (per ppm-day)	Predicted if the Model were True	100% x Ratio: Predicted / Observed	95% CI on Predicted if the Model were True
11. USEPA – Loglinear Spline – 15-yr lag (MLE) – USEPA Table 4-4 Knot @ 1,600 ppm-days	**4.89E-04**	**84.59**	**159.6%**	**(64.7, 112.9)**
12. USEPA – Loglinear Spline – 15-yr lag (95% UCL) – USEPA Table 4-4 Knot @ 1,600 ppm-days	**9.08E-04**	**141.97**	**267.9%**	**(108.5, 189.5)**
13. USEPA – Linear Spline – 15-yr lag (MLE) – USEPA Table 4-4 Knot @ 1,600 ppm-days	**7.58E-04**	**86.39**	**163.0%**	**(66.0, 115.3)**
14. USEPA – Linear Spline – 15-yr lag (95% UCL) – USEPA Table 4-4 Knot @ 1,600 ppm-days	**1.80E-03**	**135.19**	**255.1%**	**(103.4, 180.5)**

[**Boldface** values indicate that the model over-prediction of lymphoid cancer deaths is statistically significant.]

[1] The models used by Sielken & Associates and EPA [appearing as an appendix in USEPA (2016)] are the same models; however, USEPA did not use all of the individual data – Steenland et al. and USEPA only used a subsample of the individual data.

[2] The best estimate and standard error of the slope are 4.74E-06 and 3.35E-06, respectively.

[3] The 95% upper confidence limit on the slope is 1.03E-05 (4.74E-06 + 1.645×3.35E-06).

[4] The best estimate and standard error of the slope are 1.23E-05 and 2.12E-05, respectively. The standard error (2.12E-05) of the slopes was inferred from the upper bound on the slope (4.75E-05) given in Table D-36; that is 1.23E-0-5 = (4.71E-05 – 1.23E-05)/1.645.

[5] The 95% upper confidence limit on the slope is 4.71E-05 from Table D-36.

[6] The best estimate and standard error of the slope below the knot are 4.89E-04 and 2.55E-04, respectively. The slope and corresponding standard error after the knot are -4.86E-04 and 2.56E-04, respectively, from Tables 4-4 and D-33 log-linear with knot @ 1600 ppm-days.

[7] The slope after the knot is for the 95% upper confidence limit for the model is -9.07E-04 (-4.86E-04 - 1.645×2.56E-04, which a 95% LCL on the slope above the knot). This conservatively assumes perfect negative correlation of the slope below and after the knot. **Thus, the over-prediction may be larger than what is shown in the table.** The assumption of perfect negative correlation is consistent with the covariance values obtained by EPA for two-piece linear spline model; e.g., see footnote to Table D-36 in the appendices of EPA's report.

[8] The best estimate and standard error of the slope below the knot are 7.58E-04 and 6.32E-04, respectively. The slope and corresponding standard error after the knot are -7.48E-04 and 6.31E-04, respectively, from footnote to Table D-36.

[9] The slope after the knot is for the 95% upper confidence limit for the model is -1.79E-03 (-7.48E-04 - 1.645×6.32E-04, which a 95% LCL on the slope above the knot). This conservatively assumes perfect negative correlation of the slope below and after the knot. **Thus, the over-prediction may be larger than what is shown in the table.** The assumption of perfect negative correlation is consistent with the covariance values obtained by EPA (see footnote to Table D-36 in the appendices of EPA's report where the covariance is approximately equal to the negative of the variances for the slopes above and below the knot; i.e., covariance=-3.99E-07, Var1=3.99E-07, and Var2=3.98E-07).

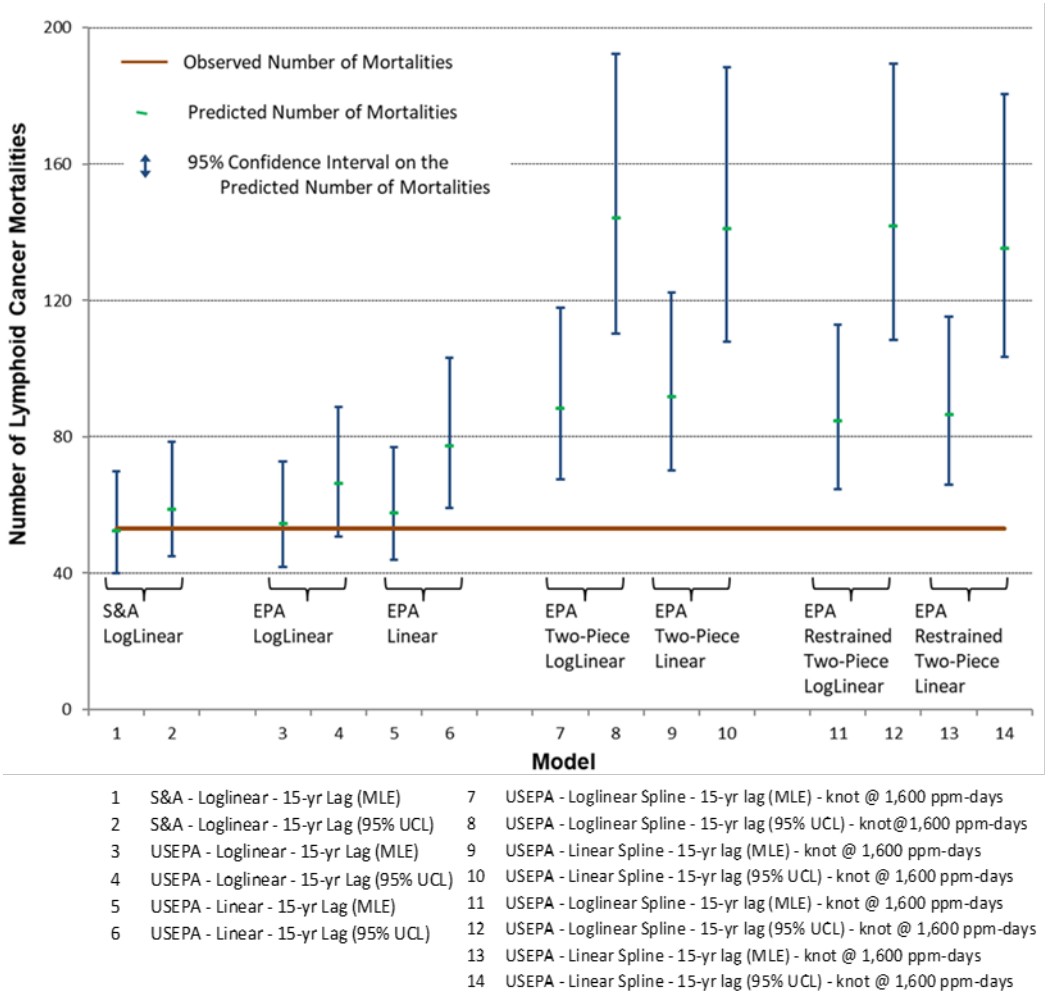


Figure 13: USEPA's Selected Model Statistically Significantly Over-Predicts Lymphoid Cancer Mortalities

A2.2 Predicted Versus Observed Number of Lymphoid Cancer Deaths in the NIOSH Cohort by Quintiles

Table 32 expands the results in Table 31 to calculate the observed and expected number of lymphoid cancer deaths in each of five quintiles. The NIOSH cohort was divided into five exposure quintiles. A total of 53 lymphoid cancer deaths were observed in the NIOSH cohort. The first quintile included the nine NIOSH workers who died with lymphoid cancer and whose cumulative exposure to EtO (lagged 15 years) was equal to zero. Cumulative exposures to EtO lagged 15 years were defined so that quintiles 2 to 5 included the same number of lymphoid cancer deaths (11) in each quintile.

Only the best estimates of the log-linear (Cox proportional hazards) model (models 1 and 3), the linear model (model 5), and the 95% upper confidence limit of the log-linear (Cox proportional hazards) model (model 2; TCEQ's selected model) predict a number of lymphoid cancer mortalities that are consistent with the number of observed deaths in each of five quintiles. USEPA's 95% UCL of the log-linear (model 4) and linear model (model 6) statistically significantly over-predict the number of the lymphoid cancer deaths in the highest exposure group.

USEPA's two-piece spline models (both the fitted models 7-10 and the restrained models 11-14) significantly over-predict the number of observed lymphoid cancer deaths at the lowest exposure quintile. *The 95% UCL of the two-piece spline models (for both the fitted models and the restrained models - models 8, 10, 12, and 14) significantly over-predict the number of observed lymphoid cancer deaths at every exposure quintile*. **More specifically, the model assessment selected by USEPA (i.e., the upper bound of the linear two-piece spline model with the "knot" at 1,600 ppm-days; 15-year exposure lag) statistically significantly over-predicts lymphoid cancer deaths for every quintile, even if the slope of the upper spline is set to zero (see Table 32 results for models 10 and 14).** The best estimates of the two-piece spline models (for both the fitted models and the restrained models - models 7, 9, 11, and 13) significantly over-predict the number of observed lymphoid cancer deaths in exposure quintiles 2 and 4 (model 9 also significantly over-predicts quintile 5).

Thus, in addition to USEPA's selected model assessment (i.e., upper bound of the linear two-piece spline model with the "knot" at 1,600 ppm-days; 15-year exposure lag) statistically significantly over-estimating the total number of observed lymphoid cancer deaths for the NIOSH cohort (141 predicted versus 53 actually observed; Table 31), their selected model also statistically significantly over-predicts lymphoid cancer deaths for every cumulative exposure group, even if the slope of the upper spline is set to zero (Table 32). The MLE of USEPA's model also statistically significantly over-predicts for every exposure quintile except quintile 3. By contrast, the model assessment selected by the TCEQ (i.e., upper bound of the log-linear/Cox proportional hazards model; 15-year exposure lag) is reasonably accurate, neither significantly over- or under-estimating lymphoid cancer deaths for cumulative exposure groups or for the cohort as a whole (59 predicted versus 53 observed).

Table 32: USEPA's Selected Model Statistically Significantly Over-Predicts Lymphoid Cancer Mortalities for All Cumulative Exposure Groups

Model [1]	Quintile 2*	Quintile 3	Quintile 4	Quintile 5
Observed	11	11	11	11
Background (No Model)	14.4 (8.0, 28.9)	7.9 (4.4, 15.9)	9.1 (5.1, 18.3)	7.4 (4.2, 14.9)
1. S&A – Loglinear – 15-yr lag (MLE)	14.4 (8.1, 28.9)	8.0 (4.5, 16.1)	9.4 (5.2, 18.8)	9.1 (5.1, 18.3)
2. S&A – Loglinear – 15-yr lag (95% UCL) - **TCEQ Adopted**	14.5 (8.1, 29.0)	8.1 (4.5, 16.2)	9.8 (5.5, 19.6)	15.0 (8.4, 30.0)
3. USEPA - Loglinear - 15-yr Lag (MLE) USEPA Table 4-2	14.4 (8.1, 29.0)	8.0 (4.5, 16.1)	9.5 (5.3, 19.1)	11.0 (6.2, 22.1)
4. USEPA - Loglinear - 15-yr Lag (95% UCL) USEPA Table 4-2	14.5 (8.1, 29.1)	8.2 (4.6, 16.4)	10.0 (5.6, 20.1)	**22.2 (12.4, 44.6)**
5. USEPA - Linear - 15-yr Lag (MLE) USEPA Table D-36	14.5 (8.1, 29.1)	8.2 (4.6, 16.5)	10.2 (5.7, 20.4)	13.2 (7.4, 26.5)
6. USEPA - Linear - 15-yr Lag (95% UCL) USEPA Table D-36	14.8 (8.3, 29.7)	9.0 (5.0, 18.0)	13.1 (7.3, 26.3)	**28.9 (16.2, 58.0)**
EPA Spline Model with Knot at 1,600 ppm-days				
7. USEPA – Loglinear Spline – 15-yr lag (MLE) – USEPA Table 4-4 Knot @ 1,600 ppm-days	**19.8 (11.1, 39.7)**	17.3 (9.7, 34.7)	**20.3 (11.3, 40.7)**	19.4 (10.8, 38.9)
8. USEPA – Loglinear Spline – 15-yr lag (95% UCL) – USEPA Table 4-4 Knot @ 1,600 ppm-days	**27.0 (15.1, 54.2)**	**33.5 (18.7, 67.3)**	**38.8 (21.7, 77.9)**	**33.3 (18.6, 66.7)**
9. USEPA – Linear Spline – 15-yr lag (MLE) – USEPA Table 4-4 Knot @ 1,600 ppm-days	**20.9 (11.7, 42.0)**	17.6 (9.8, 35.2)	**20.8 (11.6, 41.7)**	**20.9 (11.7, 41.9)**
10. USEPA – Linear Spline – 15-yr lag (95% UCL) – USEPA Table 4-4 Knot @ 1,600 ppm-days - USEPA Selected	**29.9 (16.7, 60.0)**	**30.5 (17.1, 61.2)**	**35.8 (20.0, 71.7)**	**33.4 (18.7, 67.1)**
Results using above USEPA two-piece spline models but assuming that slope for RR is zero after the "knot"				
11. USEPA – Loglinear Spline – 15-yr lag (MLE) – USEPA Table 4-4 Knot @ 1,600 ppm-days	**19.8 (11.1, 39.7)**	17.3 (9.6, 34.6)	**19.9 (11.1, 39.9)**	16.2 (9.0, 32.5)

Model [1]	Quintile 2*	Quintile 3	Quintile 4	Quintile 5
12. USEPA – Loglinear Spline – 15-yr lag (95% UCL) – USEPA Table 4-4 Knot @ 1,600 ppm-days	**27.0 (15.1, 54.2)**	**33.5 (18.7, 67.2)**	**38.6 (21.6, 77.4)**	**31.3 (17.5, 62.8)**
13. USEPA – Linear Spline – 15-yr lag (MLE) – USEPA Table 4-4 Knot @ 1,600 ppm-days	**20.9 (11.7, 42.0)**	17.5 (9.8, 35.0)	**20.1 (11.2, 40.3)**	16.4 (9.1, 32.8)
14. USEPA – Linear Spline – 15-yr lag (95% UCL) – USEPA Table 4-4 Knot @ 1,600 ppm-days	**29.9 (16.7, 60.0)**	**30.4 (17.0, 61.0)**	**35.0 (19.5, 70.2)**	**28.4 (15.9, 57.0)**

[**Boldface** values indicate that the model over-prediction of lymphoid cancer deaths for the quintile is statistically significant.]

[1] The models used to calculate the estimated number of lymphoid deaths are the same as those listed in in Table 31 and the footnotes to Table 31 apply here also. Except that the assumption of perfect negative correlation of the slopes before and after the knot in Models 8 and 10 (EPA's 95% UCL for the two-piece spline models) do not affect the predictions in quintile 2.

[*] Quintile 1 is the control (unexposed lagged-out) group with 9 lymphoid cancer mortalities observed and 11.5 mortalities predicted by all models with a 95% confidence interval of (6.0, 25.2), which includes the observed 9 lymphoid cancer deaths.

A2.3 Calculation of the Expected Number of Case-Specific Deaths in a Cohort Using US Background Hazard Rates

The SMR is a measure that compares the number of observed case-specific deaths in a study population with the number of case-specific deaths expected in the study population with known case-specific background death rates of a reference population. The case-specific background death rates of the reference population can adjust for calendar year, age, sex, race, and other relevant variables that may influence the case-specific death rates. The SMR is calculated using the following equation:

$$SMR = \frac{\sum_i y_{oi}}{\sum_i p_{oi} \frac{y_{ri}}{p_{ri}}}$$

where *i* is the stratum (the stratum is calendar year, age, sex, and race combination), y_{oi} is the number of observed deaths in the i-th stratum of the study group, p_{oi} is the number of observed person-years in the i-th stratum of the study group, y_{ri} is the number of deaths in the i-th stratum of the reference population, and p_{ri} is the number of person-years in the i-th stratum of the reference population.

The ratios $\frac{y_{ri}}{p_{ri}}$ are the stratum-specific mortality rates in the reference population. The SMR is then the ratio of the number of case-specific deaths in the study population ($\sum_i y_{oi}$) to the expected number of case-specific deaths in the study population ($\sum_i p_{oi}\frac{y_{ri}}{p_{ri}}$) estimated using the background case-specific death rates of the reference population. Several references have a more in-depth discussion of SMRs (e.g., Rothman 1986, Breslow and Day 1987, Checkoway, Pearce, and Crawford-Brown 1989).

The numerator in the SMR calculation is the sum of the calendar year, sex, race, and age-specific lymphoid cancer deaths in the NIOSH study ($\sum_i y_{oi}$) and is equal to the number of observed lymphoid cancer deaths. The denominator in the SMR calculation is the expected number of lymphoid cancer deaths in the NIOSH workers assuming that lymphoid was the only cause of death by using the US background lymphoid cancer mortality rates. The calendar year, sex, race, and age-specific lymphoid cancer mortality rates (y_{ri}/p_{ri}) for the US populations and the calendar year, sex, race, and age-specific person-years in the NIOSH study (p_{oi}) were used to calculate the expected number of the lymphoid cancer deaths in NIOSH workers.

An SMR greater than 1 (or 100%) implies that the number of observed deaths in the cohort is more than would be expected in a population with the same demographic characteristics as the cohort, except for potential exposures on the job. In contrast, an SMR less than 1 (or 100%) implies that the number of observed deaths in the cohort is less than would be expected in a population with the same demographic characteristics as the cohort, except for potential exposures on the job. The point estimate of the SMR cannot be used to derive statistically relevant conclusions indicating whether the observed number of deaths is greater or less than the expected number of deaths with a specific degree of confidence. Breslow and Day (1987) present the following equations that can be used to derive 100(1-α)% confidence intervals for the SMR.

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

and

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

where SMR_{LCL} is the 100(1-α/2)% lower confidence limit on the SMR, SMR_{UCL} is the 100(1-α/2)% upper confidence limit on the SMR, *Obs* is the number of observed cause-specific deaths (e.g., lymphoid cancer deaths) in the study ($i.e.,\ Obs = \sum_i y_{oi}$), *E* is the expected cause-specific deaths (e.g., lymphoid cancer deaths) derived from the reference population background rates

$\left(i.e.,\ E = \sum_i p_{oi} \frac{y_{ri}}{p_{ri}}\right)$, and $Z_{\alpha/2}$ is the 100(1- α/2)% percentile of the standard normal distribution.

The 100(1-α)% confidence interval for an SMR is given by the interval (SMR_{LCL}, SMR_{UCL}). Thus, if the SMR_{LCL} of a 100(1-α)% confidence interval is greater than 1 (or 100%), then the SMR is statistically significantly different (greater) than 1 (or 100%) implying that the number of observed cause-specific deaths (e.g., lymphoid cancer deaths) in the cohort is more than the number of expected cause-specific deaths (e.g., lymphoid cancer deaths) in the general population with similar demographics as the cohort. On the contrary, if the SMR_{UCL} of a 100(1-α)% confidence interval is less than 1 (or 100%), then the SMR is statistically significantly different (less) than 1 (or 100%) implying that the number of observed cause-specific deaths (e.g., lymphoid cancer deaths) in the cohort is less than the number of expected cause-specific deaths (e.g., lymphoid cancer deaths) in the general population with similar demographics as the cohort.

The US lymphoid cancer mortality rates used for the calculations of the expected number of lymphoid cancer deaths are given in Table 33 through Table 37.

A2.3.1 US Background Hazard Rates are Appropriate for Calculating the Expected Number of Lymphoid Cancer Deaths in the NIOSH Cohort due to Absence of a Healthy Worker Effect for Lymphoid Cancer Mortality

The models used by TCEQ were derived using internal comparisons and did not rely on the general U.S. population standard mortality rates. However, national rates can be used to predict the specific cancers in the NIOSH worker cohort. This is because: (1) the approach for calculating SMRs is well established and documented and has been used extensively by regulatory agencies and researchers to compare mortality rates in target populations to mortality rates in reference populations; and (2) importantly, the healthy worker effect is absent for the specific cancer endpoints of interest for the NIOSH cohort (e.g., lymphoid cancers, breast cancer), negating the potential need for internal comparisons for these particular endpoints.

Regarding these points, though opinions vary about using general population background rates for evaluating cause-specific mortality rates of occupational studies, *it is standard practice to use general population background rates* because there is often no scientific evaluation of the magnitude of the "healthy worker effect." *In fact, the standard methodology to evaluate the well-established and widely-accepted SMRs and SIRs use general population background rates*. In general, the healthy worker effect (if any) is cause-specific and often cannot be easily ascertained. However, Kirkeleit et al. (2013) researched the healthy worker effect in a large study of 366,114 randomly selected workers and compared the incidence of numerous

endpoints with the general population. Their findings indicate that there is a potential for the healthy worker effect for some endpoints while there is an increased incidence (i.e., an "unhealthy" worker effect) for other endpoints. *Relevant to the EtO assessment, Kirkeleit et al. (2013) did not find a healthy worker effect for lymphoid and hematopoietic cancer incidence, with SIRs and 95% confidence intervals of 0.97 (0.90, 1.03) and 1.09 (0.92, 1.27) for male and female workers, respectively*. The lack of a health worker effect was also true for breast cancer with an SIR and 95% confidence interval of 1.02 (0.95, 1.09).

Even more specific to the lymphoid cancer in the NIOSH cohort that drives the USEPA (2016) and TCEQ URFs, *the lymphoid cancer mortality rate in unexposed workers in the NIOSH study is not statistically significantly different from the mortality rate of the general U.S. population*. Footnote "*" to Table 33 indicates that for Quintile 1, the control (unexposed lagged-out) group, the 9 lymphoid cancer mortalities observed is well within the 95% confidence interval (6.0, 25.2) for all models. *That is, the 9 lymphoid cancer deaths observed in the unexposed male and female workers of the NIOSH cohort is consistent with the number of lymphoid cancer deaths in the general U.S. population* (i.e., during the same period of time after accounting for age, sex, and calendar year). Expressed in terms of SMRs, the SMR for lymphoid cancer deaths in the unexposed male and female NIOSH workers is equal to 0.78 (9/11.5) with a 95% confidence interval (CI) equal to (0.36, 1.50). The 95% CI on the SMR for unexposed workers includes the value of one, which indicates that *the mortality rate in the unexposed workers in the NIOSH study and the U.S. population mortality rate are not statistically significantly different* at the 5% significance level. Similar results are obtained for the male NIOSH workers that drive lymphoid cancer risk and upon which TCEQ's URF is based. More specifically, the SMR for lymphoid cancer deaths in the unexposed male NIOSH workers is equal to 1.03 (6/5.8) with a 95% CI of (0.38, 2.25). *Thus, the lymphoid cancer mortality rate in unexposed male workers in the NIOSH cohort, the gender that drives the URF, is not statistically significantly different than that in the U.S. population.*

In summary, these results demonstrate that *there is no healthy worker effect for this critical endpoint in this key group (i.e., male workers, who drive lymphoid cancer risk in the cohort and TCEQ's URF) or in males and female workers combined*. These results based on the NIOSH cohort are consistent with the findings of Kirkeleit et al. (2013) (e.g., SMRs for workers in the NIOSH cohort parallel the findings of the SIRs reported by Kirkeleit et al. for lymphoid cancer).

A2.3.2 Sensitivity Analysis Assuming a Healthy Worker Effect for Lymphoid Cancer Mortality

Despite: (1) That *the mortality rate in the unexposed workers in the NIOSH study and the U.S. population mortality rate are not statistically significantly different* (as discussed above); and (2) *The lack of a healthy worker effect for more EtO-relevant lymphoid and hematopoietic cancer* based on data from Kirkeleit et al. (as discussed above), the TCEQ conducted a sensitivity

analysis *assuming a healthy worker effect* for cancer mortality. Kirkeleit et al. (2013) indicates that the healthy worker effect for mortality from cancer is minimal, if any, overall. More specifically, Kirkeleit et al. (2013) estimates an overall cancer SMR of 0.85 and 0.84 for male and female workers, respectively. For purposes of a sensitivity analysis, the TCEQ assumed that these overall cancer SMRs apply to lymphoid cancers. That is, despite data to the contrary, the TCEQ sensitivity analysis assumes NIOSH workers were "healthier" than the general population as to cancer mortality by multiplying the U.S. male and female background hazard rates by 0.85 and 0.84, respectively, to account for the assumed healthy worker effect. The results did not change significantly. ***The MLE of USEPA's selected model still statistically significantly overestimates the number of observed (53) lymphoid deaths in the NIOSH study; 77.5 with a 95% CI of (59.3, 103.6)*. By contrast, the standard Cox proportional hazards model still estimates the observed number (53) of lymphoid deaths in the NIOSH study with a 95% confidence; 44.3 with a 95% CI of (33.9, 59.2). *Thus, even conservatively assuming a healthy worker effect (in the face of more study-specific data to the contrary), USEPA's selected model significantly overestimates the observed data it purportedly modeled*.**

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

The SMR is the ratio of observed to expected number of deaths in a cohort. The expected number of deaths is calculated assuming that the hazard rate is the background hazard rate of the reference population. However, if the background hazard rate is assumed to be affected by exposure to a carcinogen via a multiplicative function, then the expected number of deaths can be calculated assuming that the hazard rate is the product of the background hazard rate of the reference population multiplied by the exposure-response function that modifies the background rates. That is, the expected number of cause-specific deaths in a cohort can be calculated as:

$$\sum_i p_{oi} \times RR(d_i) \times \frac{y_{ri}}{p_{ri}}$$

where p_{oi} is the number of observed person-years in the i-th stratum of the study group, y_{ri} is the number of observed deaths in the i-th stratum of the reference population, p_{ri} is the number of person-years in the i-th stratum of the reference population, and $RR(d_i)$ is the exposure-response function (rate ratio function) evaluated at cumulative exposure d_i.

Using this expected number of cause-specific deaths in a cohort, an SMR* and bounds on the SMR* can be calculated as follows:

$$SMR^* = \frac{\sum_i y_{oi}}{\sum_i p_{oi} \times RR(d_i) \times \frac{y_{ri}}{p_{ri}}}$$

Similarly, the lower and upper limits of the 100(1-α)% confidence interval can be calculated as follows:

$$SMR^*_{LCL} = \frac{Obs}{E^*} \times \left(1 - \frac{1}{9 \times Obs} - \frac{Z_{\alpha/2}}{3 \times \sqrt{Obs}}\right)^3$$

and

$$SMR^*_{UCL} = \frac{(Obs + 1)}{E^*} \times \left(1 - \frac{1}{9 \times (Obs + 1)} + \frac{Z_{\alpha/2}}{3 \times \sqrt{Obs + 1}}\right)^3$$

where SMR^*_{LCL} is the 100(1-α/2)% lower confidence limit on the SMR*, SMR^*_{UCL} is the 100(1-α/2)% upper confidence limit on the SMR*, *Obs* is the number of observed cause-specific deaths (e.g., lymphoid cancer deaths) in the study $(i.e.,\ Obs = \sum_i y_{oi})$, *E** is the expected cause-specific deaths (e.g., lymphoid cancer deaths) derived from the reference population background rates multiplied by the exposure response function *RR*(d_i) $\left(i.e.,\ E^* = \sum_i p_{oi} \times RR(d_i) \times \frac{y_{ri}}{p_{ri}}\right)$, and $Z_{\alpha/2}$ is the 100(1- α/2)% percentile of the standard normal distribution.

A2.5 References

Breslow, N. E. and N. E. Day (1987). Statistical Methods in Cancer Research. Volume II – The Design and Analysis of Cohort Studies. International Agency for Research on Cancer, Lyon, France.

Checkoway, H., N. E. Pearce, and D. J. Crawford-Brown (1989). Research Methods in Epidemiology. Oxford University Press, New York, USA

Rothman, K. J. (1986). Modern Epidemiology. Little, Brown and Company, Boston USA / Toronto Canada

Table 33: Lymphoid Cancer Mortality Rates in the U.S. Population for Each Calendar Year (1930-1972), Each Race, Each Sex and Each Age Group (number of lymphoid cancer deaths per 100,000)

Age Group (Years)	1930	1940	1950	1960	1968	1969	1970	1971	1972
White Males									
< 1	0.571574	0.571574	0.571574	0.952897	0.664582	0.193834	0.250050	0.264904	0.436483
1-4	0.889715	0.889715	0.889715	0.905855	2.716523	2.469136	2.639159	2.639196	1.416049
5-9	0.896007	0.896007	0.896007	0.792474	3.181767	3.222868	3.486584	3.365958	3.053435
10-14	0.808974	0.808974	0.808974	0.764426	1.743532	2.089818	1.892907	1.777729	1.573083
15-19	1.173753	1.173753	1.173753	1.302018	2.187854	2.304943	2.062410	1.853147	1.868520
20-24	0.779566	0.779566	0.779566	1.226909	1.853888	1.437771	2.074683	1.564349	1.969677
25-34	1.246367	1.246367	1.246367	1.348092	1.948938	1.826095	1.642713	1.866738	1.436086
35-44	2.822822	2.822822	2.822822	3.369977	4.096598	4.063587	3.427241	3.219945	3.996754
45-54	6.291235	6.291235	6.291235	8.459325	10.379543	10.326954	10.435895	10.292100	9.491327
55-64	13.704865	13.704865	13.704865	18.845992	25.093104	24.651811	25.357608	27.116973	25.569775
65-74	18.092659	18.092659	18.092659	32.706133	53.237410	51.595092	51.896786	51.955307	51.216641
75-84	18.992015	18.992015	18.992015	38.781214	82.331839	88.898757	86.483903	88.585069	91.555937
85+	11.917858	11.917858	11.917858	37.471858	104.761905	101.686747	87.071343	105.399568	117.052632
Other Race Males									
< 1	0.493869	0.493869	0.493869	0.000000	0.342912	0.334609	0.950275	0.958681	1.354541
1-4	0.506669	0.506669	0.506669	0.510781	1.218451	1.163832	1.553219	0.925069	0.722674
5-9	0.875629	0.875629	0.875629	0.460755	1.440733	1.962067	1.107201	1.724138	1.617251
10-14	0.419074	0.419074	0.419074	0.374631	1.760325	1.713909	1.412963	0.949367	1.501877
15-19	0.639471	0.639471	0.639471	0.878770	2.205882	1.334380	1.415189	1.505376	1.782042
20-24	1.159879	1.159879	1.159879	0.798062	2.016607	1.771872	1.024119	1.309635	0.886525
25-34	1.371643	1.371643	1.371643	1.371711	1.282051	1.747997	1.386486	1.828030	1.277139
35-44	2.362183	2.362183	2.362183	3.357051	3.718674	3.658537	4.072298	4.099678	5.229794
45-54	5.984989	5.984989	5.984989	9.095071	11.770245	10.925926	12.172295	10.151380	12.971078
55-64	11.279807	11.279807	11.279807	17.047913	29.750000	31.365314	28.395850	31.578947	26.004728
65-74	11.984811	11.984811	11.984811	22.473431	45.908184	51.185771	46.782908	52.000000	43.314501
75-84	11.892728	11.892728	11.892728	23.349211	61.827957	62.765957	67.857013	57.692308	68.202765
85+	0.000000	0.000000	0.000000	15.943369	58.536585	52.272727	59.543142	80.851064	63.829787
White Females									
< 1	0.372830	0.372830	0.372830	0.466696	0.703416	0.752196	0.595918	0.419701	0.461215
1-4	0.589370	0.589370	0.589370	0.382623	2.033672	1.985371	1.976859	1.656868	1.449532
5-9	0.369624	0.369624	0.369624	0.240952	2.059308	2.331391	2.528940	2.320938	1.828012
10-14	0.231579	0.231579	0.231579	0.417692	1.185724	1.195589	1.110161	1.276644	1.255995
15-19	0.258359	0.258359	0.258359	0.242587	0.965624	0.882056	1.138742	1.116447	1.150775

Age Group (Years)	1930	1940	1950	1960	1968	1969	1970	1971	1972
20-24	0.521598	0.521598	0.521598	0.538865	0.859182	0.643897	0.830949	0.817682	0.823469
25-34	0.792567	0.792567	0.792567	0.695775	0.815707	0.811284	0.990505	0.730055	1.008598
35-44	1.656499	1.656499	1.656499	2.209093	2.610084	2.225193	2.125844	2.257623	2.227040
45-54	3.927054	3.927054	3.927054	5.317963	7.310358	6.770297	6.805298	6.449242	6.650224
55-64	9.581633	9.581633	9.581633	13.184796	16.236934	16.778907	16.683520	16.793724	15.473466
65-74	13.471141	13.471141	13.471141	21.389945	33.714562	34.345683	35.204790	33.589547	36.741455
75-84	13.544646	13.544646	13.544646	28.303572	54.802432	54.652880	56.864558	57.238122	56.749460
85+	11.466575	11.466575	11.466575	23.163091	57.645467	65.772669	57.425086	62.057522	59.322034
Other Race Females									
< 1	0.490851	0.490851	0.490851	0.649642	0.000000	0.343348	0.327084	0.659039	0.695476
1-4	0.255302	0.255302	0.255302	0.425917	0.788782	1.171171	1.564646	1.022305	0.545455
5-9	0.373279	0.373279	0.373279	0.153607	0.524246	0.721311	1.050270	1.136364	0.814664
10-14	0.000000	0.000000	0.000000	0.281193	1.222826	0.991408	0.837986	1.144310	0.629327
15-19	0.302773	0.302773	0.302773	0.122783	0.642055	1.078582	0.663027	0.921986	0.679348
20-24	0.572140	0.572140	0.572140	0.142154	1.020408	0.287632	0.898678	0.583333	0.960769
25-34	0.686160	0.686160	0.686160	0.906197	1.654997	1.175015	0.652594	0.694444	0.986842
35-44	1.574455	1.574455	1.574455	3.092078	2.105978	2.642276	2.321355	2.675585	2.514891
45-54	4.516905	4.516905	4.516905	7.099807	9.083333	9.046455	8.699902	8.268934	8.308157
55-64	7.848951	7.848951	7.848951	10.717328	20.000000	16.902944	18.750576	20.582121	16.276704
65-74	5.746153	5.746153	5.746153	12.368748	30.629139	27.597403	28.920872	31.981279	33.027523
75-84	4.880954	4.880954	4.880954	16.111612	37.500000	33.333333	32.715935	35.000000	34.437086
85+	0.000000	0.000000	0.000000	12.414341	29.508197	33.846154	22.881259	42.465753	36.842105

Table 34: Lymphoid Cancer Mortality Rates in the U.S. Population for Each Calendar Year (1973-1981), Each Race, Each Sex and Each Age Group (number of lymphoid cancer deaths per 100,000)

Age Group (Years)	1973	1974	1975	1976	1977	1978	1979	1980	1981
White Males									
< 1	0.908058	0.224475	0.528294	0.300067	0.500615	0.358533	0.273877	0.132507	0.132064
1-4	2.244898	1.937849	1.833031	1.491692	1.211771	1.370124	1.234337	0.999559	1.346066
5-9	3.192572	3.142184	2.786254	3.041926	2.701618	2.013605	2.703456	2.514574	2.153795
10-14	2.131166	2.046687	1.720841	1.787372	2.181993	1.920932	1.734473	1.758458	1.563759
15-19	1.934907	1.908439	1.957140	1.817788	1.691974	1.677743	1.720171	1.719677	1.542872
20-24	1.456249	1.256932	1.508621	1.205242	1.383173	1.537081	1.481645	1.646638	1.395948
25-34	1.559640	1.639344	1.467136	1.432200	1.456079	1.578878	1.322802	1.543315	1.499603

Age Group (Years)	1973	1974	1975	1976	1977	1978	1979	1980	1981
35-44	3.285860	3.206107	3.239279	2.932876	2.984485	3.414495	3.156437	3.505926	3.005275
45-54	9.415647	10.002913	9.567420	9.625196	9.086395	9.480337	9.692479	9.433185	9.489925
55-64	24.776732	24.812299	25.402042	24.272853	24.671202	24.745497	24.588897	25.549930	25.109082
65-74	52.533589	52.720450	50.549249	52.758868	52.749171	53.199113	54.677339	54.513390	52.882396
75-84	91.595563	91.298812	90.050167	92.269737	90.846216	96.881248	98.868072	98.827567	99.726331
85+	109.183673	109.126214	119.074074	116.333938	119.789842	125.252525	135.008104	135.478217	128.314866
Other Race Males									
< 1	0.000000	0.350064	0.000000	0.686344	0.000000	0.952922	0.604677	0.000000	0.000000
1-4	0.890472	1.334520	1.432408	1.648352	0.925926	0.915751	0.896057	0.867085	1.145101
5-9	1.717033	1.670146	1.742160	1.098901	2.105978	1.683502	1.346801	0.799939	1.551788
10-14	1.607916	1.411909	0.973828	1.039755	1.363918	1.322418	0.890019	1.453699	1.239236
15-19	1.851852	1.726343	1.179392	1.390568	1.014925	1.410106	1.567034	1.377656	1.363956
20-24	1.528014	1.383238	1.242236	1.187825	1.275691	1.709986	1.058901	1.480282	1.175116
25-34	1.333333	1.145475	1.243243	1.379663	1.699854	1.661283	1.179554	1.310302	1.284428
35-44	3.903201	2.773498	3.506098	3.048327	3.537906	3.778866	3.653586	3.462009	4.639626
45-54	9.490940	13.356164	10.365336	10.867734	10.067114	9.468439	11.367381	10.689003	10.210284
55-64	27.570093	29.633867	29.319955	30.363036	28.862661	25.991649	29.183673	29.668996	26.891935
65-74	56.880734	54.821429	53.739130	53.962901	54.545455	58.582677	50.844854	58.720972	54.042417
75-84	73.991031	76.855895	66.115702	74.806202	81.992337	76.226415	78.651685	85.585907	93.874677
85+	64.583333	76.000000	75.925926	60.000000	82.142857	108.620690	106.779661	80.643834	104.987699
White Females									
< 1	0.559929	0.396269	0.479311	0.555150	0.302594	0.455050	0.361702	0.210232	0.139542
1-4	1.087926	1.337486	1.087164	1.130952	1.031553	1.022044	0.964947	0.643648	0.888346
5-9	2.089711	1.931242	1.779013	1.525870	1.558551	1.671667	1.377491	1.181182	1.282891
10-14	1.010913	1.042753	0.977275	0.935829	1.054746	0.896104	0.828655	0.922761	1.031858
15-19	1.049838	0.888990	0.972081	0.705803	0.887341	0.700328	0.797176	0.818234	0.945110
20-24	0.683717	0.843359	0.774256	0.900794	0.672464	0.716642	0.628578	0.724198	0.705556
25-34	0.861660	0.811775	0.928295	0.739332	0.837019	0.936504	0.798198	0.855556	0.724416
35-44	2.267551	2.112676	2.106728	1.792044	1.865996	1.696495	1.630139	1.887533	1.727053
45-54	6.246017	6.551095	6.287809	6.452209	6.487905	6.471816	6.256618	6.115654	5.936539
55-64	16.013353	16.622439	15.990803	16.423433	16.627989	16.348638	16.209867	16.803601	17.030421
65-74	34.125587	34.821812	32.178287	34.755847	34.549814	35.034501	35.199592	37.603777	35.889455
75-84	58.124174	58.643892	57.581864	61.363079	61.298077	61.771617	63.731992	67.535625	68.589388
85+	67.239636	66.761364	67.724868	67.617450	76.367962	76.519130	75.692964	84.172570	83.353422
Other Race Females									
< 1	0.718184	0.000000	0.000000	0.000000	0.000000	0.654986	0.311744	0.000000	0.000000
1-4	0.898473	0.450045	1.364877	0.372439	0.753296	0.279851	0.547445	0.795146	0.583260
5-9	0.966851	0.629811	1.190476	0.968188	0.959561	0.886767	0.752394	0.407426	1.169315

Age Group (Years)	1973	1974	1975	1976	1977	1978	1979	1980	1981
10-14	0.623053	0.992556	0.802965	0.745805	0.693569	0.960307	0.774693	0.642377	0.757866
15-19	0.786885	0.571429	0.803461	0.422705	0.774732	0.587544	0.815376	0.864307	0.402981
20-24	0.538462	0.591716	0.283487	0.683060	0.654879	0.758534	0.612745	0.654753	0.634340
25-34	0.677083	0.935961	0.836431	0.924296	0.962343	0.558659	0.833018	1.034294	0.828562
35-44	2.156863	2.450032	1.977041	2.114428	2.238355	2.231356	2.103468	2.399917	2.864034
45-54	9.830007	6.540698	9.305655	6.770099	8.432056	6.662088	8.316430	8.035665	6.734315
55-64	18.818819	17.543860	19.038643	20.702403	19.516562	20.555074	18.891688	19.739761	18.660537
65-74	37.037037	34.240688	32.088520	34.087883	32.101911	32.885086	35.924617	32.425347	40.174421
75-84	31.761006	36.445783	44.067797	45.212766	48.041775	45.641026	47.727273	57.289609	57.167055
85+	46.250000	54.117647	41.935484	43.877551	45.192308	50.000000	63.157895	65.743449	70.517392

Table 35: Lymphoid Cancer Mortality Rates in the U.S. Population for Each Calendar Year (1982-1990), Each Race, Each Sex and Each Age Group (number of lymphoid cancer deaths per 100,000)

Age Group (Years)	1982	1983	1984	1985	1986	1987	1988	1989	1990
White Males									
< 1	0.000000	0.462407	0.000000	0.192266	0.064567	0.512302	0.000000	0.244261	0.118477
1-4	0.897367	1.310122	0.781290	0.830986	0.877404	0.739505	0.737235	0.663349	0.708275
5-9	2.366171	1.846937	1.510829	1.428039	1.366221	1.467699	1.225459	1.297239	0.913484
10-14	1.583212	1.360994	1.426616	1.285190	1.274476	1.210121	1.201909	1.428199	1.352777
15-19	1.796605	1.780555	1.689925	1.682906	1.512290	1.333880	1.353366	1.212178	1.409300
20-24	1.343823	1.284539	1.270779	1.324499	1.419361	1.497749	1.274751	1.514134	1.248516
25-34	1.527609	1.570647	1.584635	1.706365	2.154965	1.607166	1.992268	1.977337	2.268786
35-44	3.607424	3.210907	3.607591	3.900018	3.907493	3.733309	3.744332	4.073447	3.925666
45-54	10.320582	9.492029	9.475140	9.981628	10.353269	10.305775	10.121232	10.454357	11.342008
55-64	25.740401	25.933995	26.359149	27.642635	26.093181	28.162326	28.577168	29.628210	29.421239
65-74	55.446249	58.683266	58.006916	60.547081	63.379973	61.768858	60.894609	63.835855	64.680548
75-84	102.512985	103.269530	102.903810	113.797884	111.957418	110.325657	117.539257	121.572182	124.689270
85+	141.091466	154.657919	146.182157	158.545624	152.478016	146.762825	171.258407	163.709977	185.700410
Other Race Males									
< 1	0.282407	0.000000	0.560626	0.544009	0.265887	0.513383	0.243094	0.231537	0.000000
1-4	0.950552	0.843139	0.898864	0.815968	0.584038	0.359246	0.352241	0.545662	0.529965
5-9	1.544365	1.263091	1.035059	1.065461	1.635687	1.002256	0.802618	0.847424	0.838924
10-14	1.101152	1.094825	1.341328	1.465289	1.305275	0.991744	0.674730	1.075256	0.990555
15-19	1.544260	1.214203	1.108428	0.701977	0.978176	1.531826	1.121842	1.232062	0.892218

Age Group (Years)	1982	1983	1984	1985	1986	1987	1988	1989	1990
20-24	0.848498	1.603323	1.108261	1.322919	1.200467	0.919044	1.446631	1.389804	1.442548
25-34	1.840239	1.941467	1.637358	1.906600	1.752430	1.457848	1.865610	2.782049	2.290311
35-44	3.630473	3.495188	4.120332	4.426983	4.713920	4.554605	4.972986	4.699949	5.240313
45-54	12.753297	11.795082	11.153652	10.804774	11.090469	11.424834	12.745138	13.021074	13.059052
55-64	27.441584	33.281437	30.656579	29.982650	30.277039	26.602320	29.171684	30.098894	33.984171
65-74	57.237298	55.381074	50.838187	61.469040	67.722773	64.142203	60.374990	60.402824	65.684984
75-84	99.028610	108.712639	94.311838	97.257155	112.593187	106.228728	99.871509	110.026091	109.071026
85+	110.976140	120.734757	82.336687	113.366296	106.579982	137.074874	121.273370	148.091471	159.703198
White Females									
< 1	0.412871	0.418804	0.207705	0.338393	0.204025	0.337325	0.397082	0.450230	0.062415
1-4	0.740887	0.943464	0.464971	0.714428	0.693092	0.601971	0.653006	0.419260	0.451249
5-9	1.294763	0.911457	0.835611	0.988693	0.757493	0.627520	0.559821	0.641137	0.623382
10-14	0.811883	0.631763	0.881446	0.834117	0.803605	0.716906	0.557631	0.640258	0.556603
15-19	0.816159	0.870140	0.723414	0.626600	0.838982	0.794999	0.644126	0.647127	0.788964
20-24	0.873275	0.679190	0.641055	0.778479	0.804127	0.708784	0.656806	0.791296	0.786603
25-34	0.743563	0.696736	0.814677	0.906247	0.940198	0.770082	0.829128	0.869329	0.884170
35-44	1.741456	1.859996	2.115381	1.992830	1.956782	1.717332	2.159311	1.856792	1.787279
45-54	6.734416	6.563147	6.457907	6.609959	6.253106	6.042936	6.355324	6.076045	6.084263
55-64	16.917034	17.085084	17.960658	18.684330	17.474939	17.735989	17.586514	18.798277	17.622023
65-74	37.596194	39.177268	39.824889	39.607408	41.121751	40.965889	41.342613	43.020215	43.082987
75-84	69.543091	70.552506	72.529403	71.315776	76.337351	76.845877	77.916555	80.989763	81.092049
85+	92.412534	89.912880	93.843998	94.727554	100.448726	104.084539	103.516519	109.816269	114.634887
Other Race Females									
< 1	0.292722	0.000000	0.868817	0.563369	0.553598	0.000000	0.252484	0.239977	0.468898
1-4	0.726035	0.546679	0.611366	0.454753	0.298587	0.515052	1.010791	0.699719	0.476427
5-9	0.548698	1.087145	0.198370	0.640049	0.804902	0.421807	0.645421	0.520951	0.458591
10-14	0.812410	0.622286	0.437587	0.752269	0.382603	0.509268	0.377932	0.490451	0.477840
15-19	0.580762	0.764674	0.593717	0.298791	0.471507	0.640464	0.461812	0.519634	0.748110
20-24	0.853074	0.561540	0.501356	0.221421	0.554927	0.671071	0.564213	0.510058	0.851649
25-34	0.731149	0.674739	0.950363	1.008959	0.926506	0.903771	1.071554	0.710502	0.963634
35-44	2.213313	2.192893	2.291606	2.543862	2.321505	2.242482	2.132750	2.326151	2.652870
45-54	7.298407	7.121108	7.312326	6.550464	8.025120	7.634042	7.331957	7.589449	8.253123
55-64	18.533248	17.381368	20.156957	19.876547	18.758072	18.216235	19.695708	19.588978	19.595873
65-74	37.355813	38.276541	36.088017	38.533843	40.391660	39.156632	40.894103	41.773392	41.612207
75-84	59.725264	61.003109	58.979590	72.662063	61.616938	61.855941	67.427820	70.322620	71.910686
85+	64.834220	66.926697	64.149876	77.144586	79.929917	83.506794	81.033922	81.645237	83.769867

Table 36: Lymphoid Cancer Mortality Rates in the U.S. Population for Each Calendar Year (1991-1999), Each Race, Each Sex and Each Age Group (number of lymphoid cancer deaths per 100,000)

Age Group (Years)	1991	1992	1993	1994	1995	1996	1997	1998	1999
White Males									
< 1	0.120549	0.304542	0.309342	0.250062	0.125911	0.126229	0.381286	0.313145	0.261647
1-4	0.598010	0.634873	0.641730	0.483114	0.597917	0.525628	0.322071	0.389179	0.520896
5-9	1.077332	1.046375	0.842215	0.869082	1.071523	0.627185	0.728541	0.635617	0.535847
10-14	1.069727	0.922609	1.018617	0.953443	0.855020	0.884591	0.804178	0.847763	0.589373
15-19	1.394160	1.411226	1.281312	1.131257	1.049657	1.046720	0.934061	1.187142	0.880738
20-24	1.486628	1.485252	1.049435	1.532901	1.098601	1.291260	1.508268	1.552742	1.398208
25-34	2.153514	2.230164	2.090814	2.252798	2.244475	2.011220	2.201578	1.773869	1.305571
35-44	4.716193	4.434700	4.386889	4.381832	4.635446	4.322717	3.891075	3.694620	2.936410
45-54	11.299132	10.765887	10.498471	11.240728	10.956518	10.384872	10.941259	10.085568	9.264970
55-64	28.990578	28.964490	28.869688	30.789233	30.267561	29.977605	29.599598	28.278056	27.768360
65-74	65.820142	67.437957	67.622686	70.574494	70.831434	69.983251	72.455585	71.013446	69.063573
75-84	123.244041	128.192453	129.169255	130.541394	132.139030	135.097298	134.542905	135.014407	136.039499
85+	184.620012	182.774888	186.482519	202.084388	203.049861	205.679170	195.813850	199.761637	200.496795
Other Race Males									
< 1	0.000000	0.000000	0.231198	0.000000	0.490283	0.492542	0.242734	0.476757	0.000000
1-4	0.251040	0.180786	0.291989	0.172394	0.286071	0.287824	0.233362	0.352567	0.176170
5-9	0.706327	0.689215	0.565082	0.492402	0.520381	0.819514	0.572628	0.430521	0.256131
10-14	0.775427	0.641820	0.568414	0.759836	1.047504	0.733418	0.767420	0.561479	0.813209
15-19	1.191880	1.185346	0.500675	0.864956	1.198790	0.553187	0.731660	0.662851	1.070727
20-24	1.124612	1.642354	1.785301	1.508855	0.972847	1.313934	2.015238	0.645289	0.993891
25-34	2.237519	2.484545	2.407845	2.206208	2.567098	2.425574	2.111731	1.761624	1.717844
35-44	5.264830	5.221627	4.846035	4.669117	5.130747	5.026924	5.259584	4.383872	3.907748
45-54	12.192547	12.871079	12.740362	12.099461	12.981341	12.574332	13.039173	11.972081	9.760551
55-64	31.597492	34.051901	28.743845	34.058142	31.510938	32.051830	30.667501	30.433409	31.292855
65-74	67.516141	61.893730	69.133246	62.181494	62.604246	67.819297	64.586214	62.510594	61.446247
75-84	118.346204	108.465272	111.503892	101.134128	110.952607	117.171986	116.895856	108.432653	108.149986
85+	131.534134	140.571056	164.607271	156.009507	161.524956	154.217709	152.287127	162.763360	161.416252
White Females									
< 1	0.189610	0.128216	0.260841	0.394373	0.198615	0.463996	0.600393	0.328510	0.206611
1-4	0.544654	0.484663	0.362290	0.393668	0.231834	0.268299	0.322384	0.375495	0.411102
5-9	0.617083	0.712038	0.651712	0.505619	0.510744	0.422820	0.559046	0.412139	0.282375
10-14	0.420396	0.650159	0.510683	0.558181	0.525734	0.507201	0.530655	0.539522	0.375783
15-19	0.791386	0.689823	0.563043	0.653104	0.495588	0.564889	0.605686	0.474534	0.521361

Age Group (Years)	1991	1992	1993	1994	1995	1996	1997	1998	1999
20-24	0.719853	0.647753	0.577305	0.783432	0.732804	0.840555	0.913694	0.930414	0.701500
25-34	0.928258	0.984040	0.944766	1.037638	0.882957	1.072279	0.822517	0.832823	0.824799
35-44	1.920846	1.937426	1.865423	2.084310	2.097702	1.968226	1.983071	1.727557	1.672751
45-54	6.500862	5.997125	5.912764	6.459897	6.114375	6.139397	5.639134	5.577498	5.202266
55-64	19.178724	18.330817	19.220898	19.593339	19.239323	19.268723	19.531043	17.763069	17.363737
65-74	44.670651	45.063962	46.706389	46.334466	47.634353	46.662600	47.170072	45.873513	46.282577
75-84	85.652607	85.539274	87.768235	88.536784	89.289949	90.527655	89.550870	91.065418	91.226321
85+	118.035157	115.502420	120.620701	117.264248	125.040442	121.648591	124.871721	121.364315	122.155611
Other Race Females									
< 1	0.234086	0.000000	0.000000	0.000000	0.254598	0.254855	0.504694	0.000000	1.249619
1-4	0.193747	0.434289	0.180589	0.415097	0.472506	0.356208	0.300468	0.120879	0.181199
5-9	0.502308	0.109141	0.688359	0.355915	0.489020	0.376693	0.364674	0.178399	0.221088
10-14	0.340783	0.658581	0.265457	0.260343	0.718685	0.604677	0.148552	0.193467	0.420867
15-19	0.760147	0.290665	0.629617	0.667091	0.589240	0.516753	0.551219	0.243356	0.478619
20-24	0.552215	0.701958	0.744932	0.369962	0.529128	0.641656	0.371187	0.574389	0.811758
25-34	1.250760	1.161703	1.074879	0.969668	1.282122	1.191926	1.034714	1.221072	0.860489
35-44	2.631571	2.695297	2.201742	2.072282	2.737377	2.480527	2.904835	2.831665	2.114252
45-54	7.433460	7.524094	7.964662	7.841874	7.423539	6.577967	6.862564	6.910658	6.250333
55-64	20.877164	19.463921	21.271408	20.568934	23.617713	21.535597	20.943180	21.726642	21.037674
65-74	46.704315	41.136051	43.407193	39.603040	41.951707	46.011816	43.479905	44.474852	41.977259
75-84	81.049219	72.227947	77.173631	76.716888	75.573071	76.119672	72.954561	78.245435	76.115208
85+	87.337153	99.305842	94.501598	94.680398	94.904241	99.516750	98.701031	99.677092	95.995562

Table 37: Lymphoid Cancer Mortality Rates in the U.S. Population for Each Calendar Year (2000-2008), Each Race, Each Sex and Each Age Group (number of lymphoid cancer deaths per 100,000)

Age Group (Years)	2000	2001	2002	2003	2004	2005	2006	2007	2008
White Males									
< 1	0.524806	0.250750	0.381423	0.126342	0.125603	0.063462	0.378854	0.433816	0.375811
1-4	0.390715	0.311593	0.340849	0.547846	0.383588	0.428761	0.414535	0.207105	0.460199
5-9	0.647961	0.536133	0.544783	0.809098	0.738830	0.586288	0.440868	0.721561	0.485417
10-14	0.836564	0.644528	0.792704	0.683952	0.508571	0.705677	0.615860	0.597909	0.405742
15-19	1.143733	1.118192	1.005208	0.941732	1.015803	0.933706	0.867502	0.827787	0.838181
20-24	1.424321	1.262936	1.335348	1.160621	1.051160	1.247020	1.314343	1.043871	1.270049
25-34	1.207456	1.325997	1.292035	1.232081	1.287954	1.026088	1.180857	1.123533	1.249620

Age Group (Years)	2000	2001	2002	2003	2004	2005	2006	2007	2008
35-44	2.951331	2.947883	2.787913	2.719071	2.445056	2.470472	2.151277	2.365903	2.161794
45-54	8.736368	8.658735	8.160044	7.522465	7.274624	6.838794	6.861847	6.613099	6.164806
55-64	26.024599	25.768249	24.602045	24.337611	22.290379	21.443948	20.815903	20.218269	20.093016
65-74	68.210725	66.846157	66.754466	63.724138	59.058038	59.772839	55.443301	55.225882	52.210701
75-84	137.861646	131.603614	132.026187	129.571266	125.750437	126.843740	126.655258	125.431566	123.714919
85+	202.953378	206.959834	212.138265	213.290538	201.174047	212.220517	195.502713	202.949122	202.726728
Other Race Males									
< 1	0.235491	0.000000	0.448970	0.000000	0.000000	0.211882	0.207428	0.000000	0.389636
1-4	0.232676	0.174487	0.114159	0.281887	0.388513	0.436998	0.324330	0.529700	0.359809
5-9	0.426663	0.433151	0.350934	0.177529	0.536648	0.669715	0.307361	0.432344	0.255016
10-14	0.352086	0.844244	0.697316	0.803100	0.437740	0.359507	0.481909	0.444312	0.486827
15-19	0.920683	1.076046	0.792248	0.602980	0.459569	0.604006	0.779758	0.720078	0.890076
20-24	1.679528	1.056120	0.877657	1.167735	1.357733	1.165263	1.232959	1.051449	0.744980
25-34	1.363152	1.404313	1.538684	1.551104	1.403061	1.602819	1.098655	1.126761	1.266334
35-44	2.835120	3.817562	3.392236	3.049851	2.553021	2.602693	3.074193	3.089058	2.116457
45-54	10.717689	9.866223	8.851983	9.939288	9.058168	9.391368	8.899028	8.540407	7.925244
55-64	26.363186	29.985785	26.175855	23.212888	23.481933	23.096876	24.894886	21.742272	21.917414
65-74	61.467682	61.255497	57.822519	52.268589	57.715894	54.302768	52.212361	49.404447	51.758535
75-84	102.947245	104.276589	99.069233	95.457067	100.239504	96.713415	94.921776	97.159675	93.011377
85+	145.308316	142.557723	134.973258	143.433958	145.190271	126.514193	152.502927	143.278205	131.946501
White Females									
< 1	0.483682	0.131239	0.332853	0.596126	0.263276	0.199731	0.198550	0.324862	0.327583
1-4	0.376789	0.310412	0.392293	0.388978	0.217928	0.199665	0.334287	0.317396	0.216318
5-9	0.425186	0.446824	0.547368	0.446350	0.436685	0.356507	0.299872	0.379088	0.375590
10-14	0.486294	0.377656	0.561295	0.397890	0.411565	0.441312	0.381939	0.540134	0.375560
15-19	0.492428	0.502412	0.435949	0.420339	0.629975	0.422781	0.479903	0.488373	0.438460
20-24	0.606969	0.729405	0.791141	0.676381	0.607536	0.555826	0.530911	0.682503	0.390786
25-34	0.751260	0.854954	0.782482	0.621166	0.630221	0.725255	0.731735	0.641508	0.582598
35-44	1.522875	1.588986	1.609632	1.453520	1.243847	1.286495	1.359781	1.251519	1.204327
45-54	5.326357	4.737304	4.630905	4.389539	4.295574	3.898529	3.933733	3.694953	3.534546
55-64	17.389128	16.335271	15.009996	13.676430	13.322191	13.352400	12.130725	11.797667	11.197640
65-74	44.010466	41.752191	40.585987	37.403030	36.937724	35.289786	35.434227	33.258375	31.591145
75-84	90.119912	87.396791	84.699781	84.711257	82.164651	81.038234	78.777329	78.024018	75.235482
85+	128.513697	128.834098	129.776449	128.647982	124.750168	125.342160	126.731086	123.320293	121.223154
Other Race Females									
< 1	0.244260	0.000000	0.464279	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
1-4	0.359362	0.179663	0.176290	0.232051	0.114423	0.000000	0.000000	0.164215	0.053145
5-9	0.309062	0.402679	0.271573	0.228525	0.459707	0.000000	0.135214	0.266130	0.261604

Age Group (Years)	2000	2001	2002	2003	2004	2005	2006	2007	2008
10-14	0.227928	0.174845	0.254859	0.499492	0.206140	0.289557	0.373864	0.083534	0.377093
15-19	0.520827	0.465908	0.824630	0.536728	0.260194	0.208961	0.283326	0.236140	0.231250
20-24	0.838657	0.702065	0.770675	0.398600	0.393036	0.650290	0.687329	0.466475	0.581676
25-34	1.000629	1.272210	1.020700	0.869944	0.899656	0.752461	0.696625	0.664100	0.611427
35-44	2.317793	2.049276	1.899200	1.862371	1.737403	2.008196	1.872617	1.809375	1.348465
45-54	6.319216	6.213190	6.929462	5.666120	5.479445	5.300950	5.361658	5.400012	4.546107
55-64	17.592975	18.765077	17.788091	14.672254	15.503902	15.881942	14.640494	14.890397	13.472998
65-74	40.580024	41.223164	41.278055	41.797987	36.900825	36.086683	34.291068	34.010516	31.508649
75-84	74.119505	74.499069	70.453876	77.651645	71.641475	61.796102	62.880913	66.641937	62.963260
85+	115.616309	97.336673	86.333420	98.078476	99.450371	89.589566	92.445974	88.253258	86.059963

Appendix 3 Hypothetical Example of Appearance of Supra-Linearity in the Absence of Truly Low-Dose Data

USEPA acknowledges that "the actual exposure-response relationship at low exposure levels is unknown" (pp. 4-61 and 4-74 of USEPA 2016). The inability to observe sublinearity in the NIOSH cohort might be explained by the lack of dose-response data at low air concentrations (e.g., beginning ≈0.5 ppb) that would allow total internal exposures (endogenous + exogenous) to remain in/near the normal endogenous range (e.g., see Figure 3 and Figure 6). Where available dose-response data are predominated by exposures above the area in the dose-response expected to be sublinear (i.e., within/near/below the normal endogenous range in the present case), if the doses are sufficiently high to be above the upward inflection point, then the dose-response observed based on the data available might appear supra-linear overall. As a hypothetical example, Figure 14 below is similar to Figure 4-2 of USEPA (2016) for lymphoid cancer.

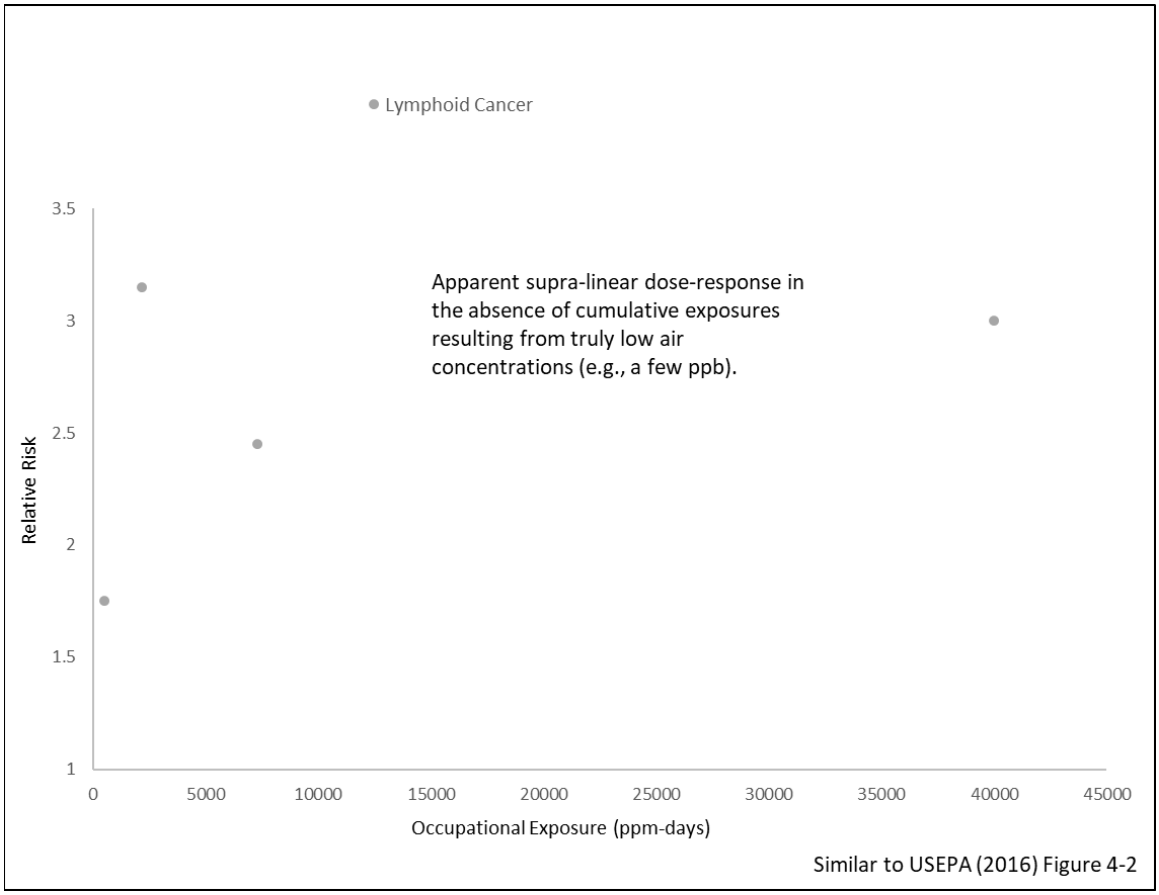


Figure 14: Seemingly Supra-linear Dose-Response for Lymphoid Cancer

The dose-response as presented (not based on the individual data or additional exposure groups) may appear overall supra-linear in nature, as noted by USEPA (2016). However, examination of the dose axis reveals that there are no truly low-dose data to characterize the shape of the dose-response at low exposures, especially within/near/below the endogenous range where both the TCEQ and USEPA would expect sublinearity (e.g., ≥0.5 ppb). Hypothetical dose-response data in the range of endogenous exposures and below were used to produce Figure 15 and Figure 16 (see below).

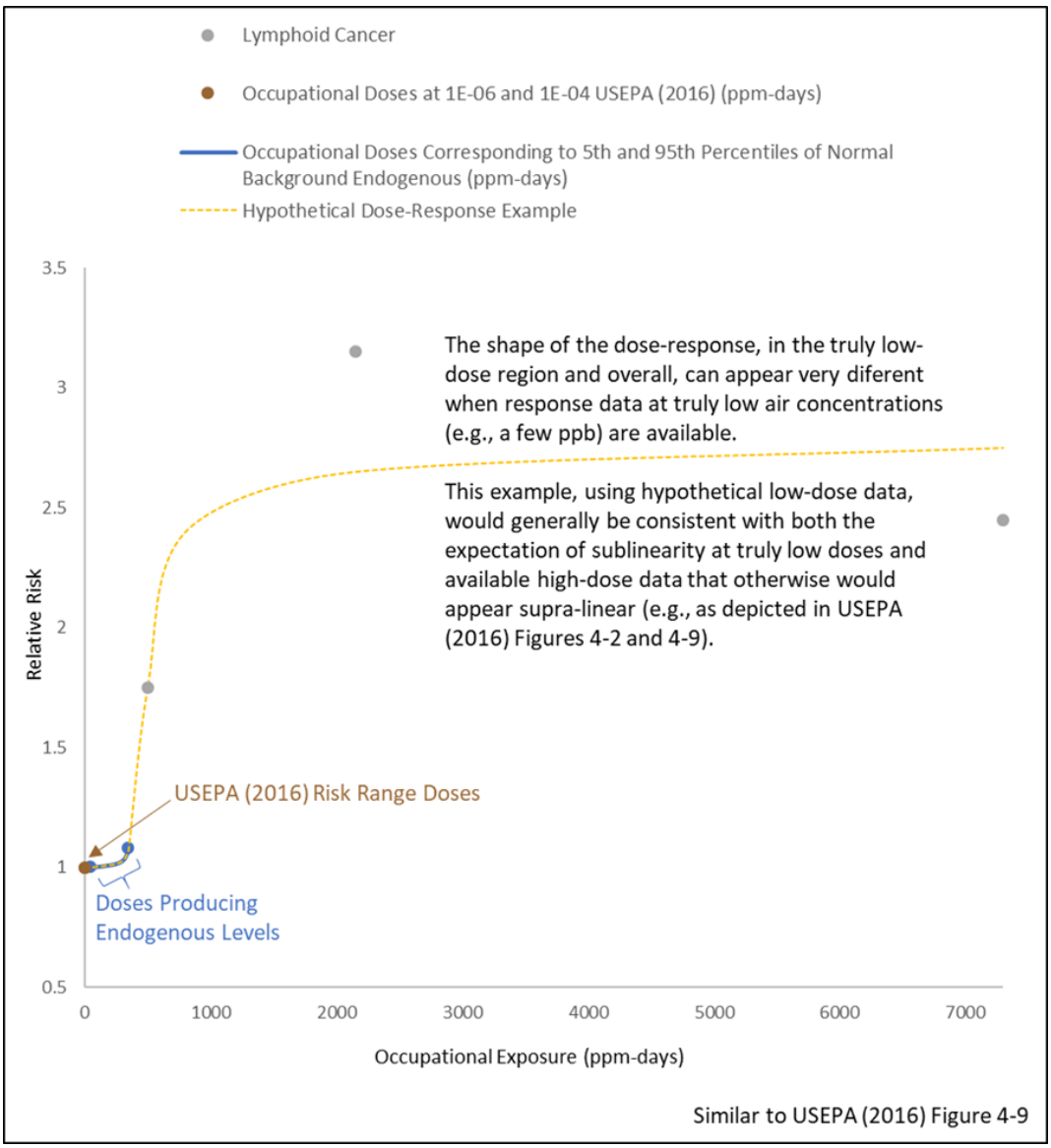


Figure 15. Hypothetical Sublinear Dose-Response at Truly Low Doses Plotted with Available High-Dose Data for Lymphoid Cancer

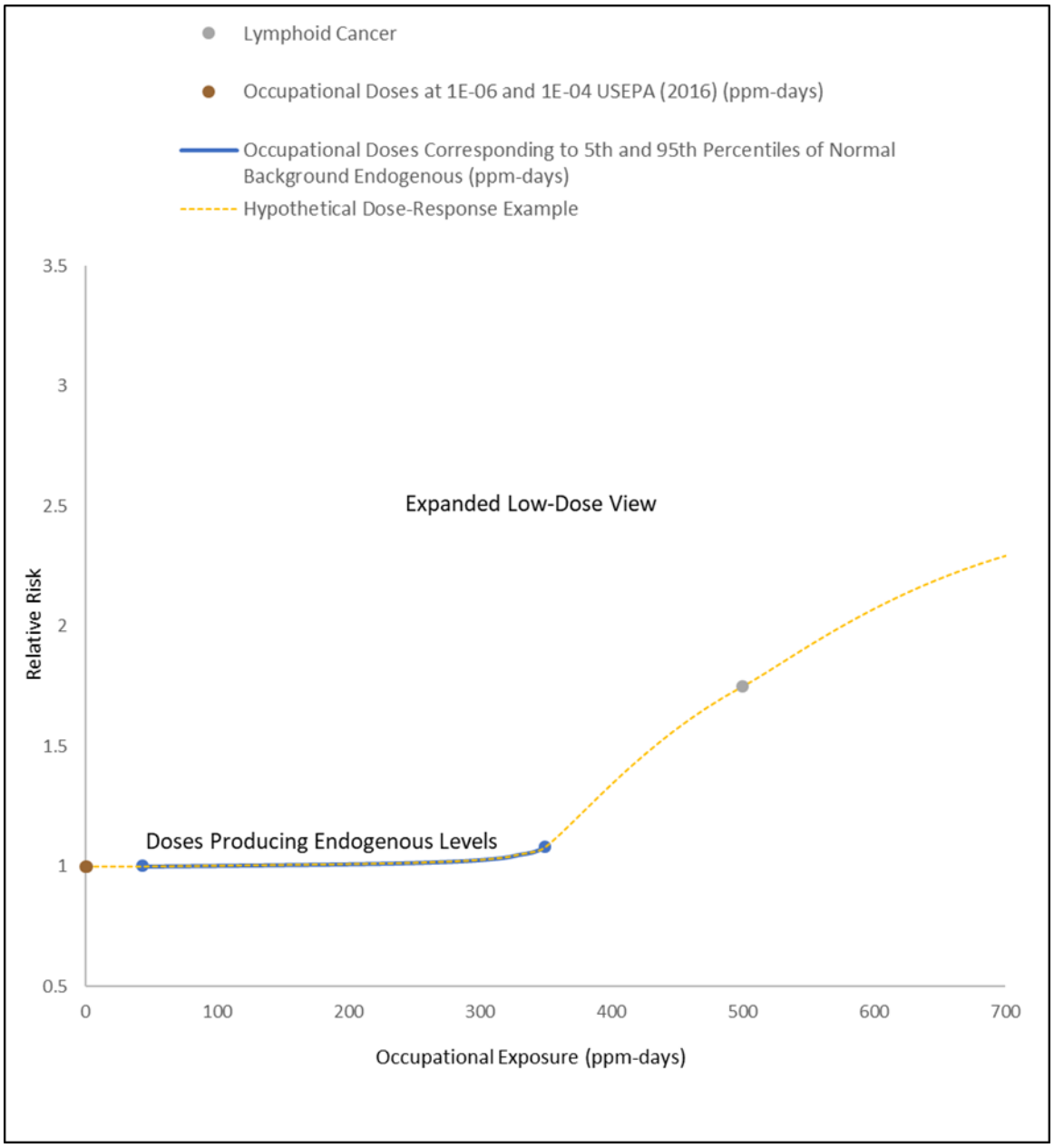


Figure 16: Hypothetical Sublinear Dose-Response at Truly Low Doses Plotted with Available High-Dose Data for Lymphoid Cancer – Expanded Low-Dose View

The availability of adequate, truly low-dose data in this hypothetical example reveals the existence of sublinearity in the overall dose-response at doses corresponding to the endogenous range (and significantly lower doses corresponding to 1E-06 to 1E-04 excess risk based on USEPA 2016). However, simple removal of these truly low-dose data results in a graph

depicting a seemingly supra-linear dose-response (Figure 17) with a steep low-dose slope down to a relative risk of 1 at 0 dose (similar to Figure 4-9 in USEPA 2016). At the same time, it should be realized that use of a different (e.g., higher) number of cumulative exposure intervals provides a different visual impression (e.g., see Figure 6S of Valdez-Flores et al. 2013).

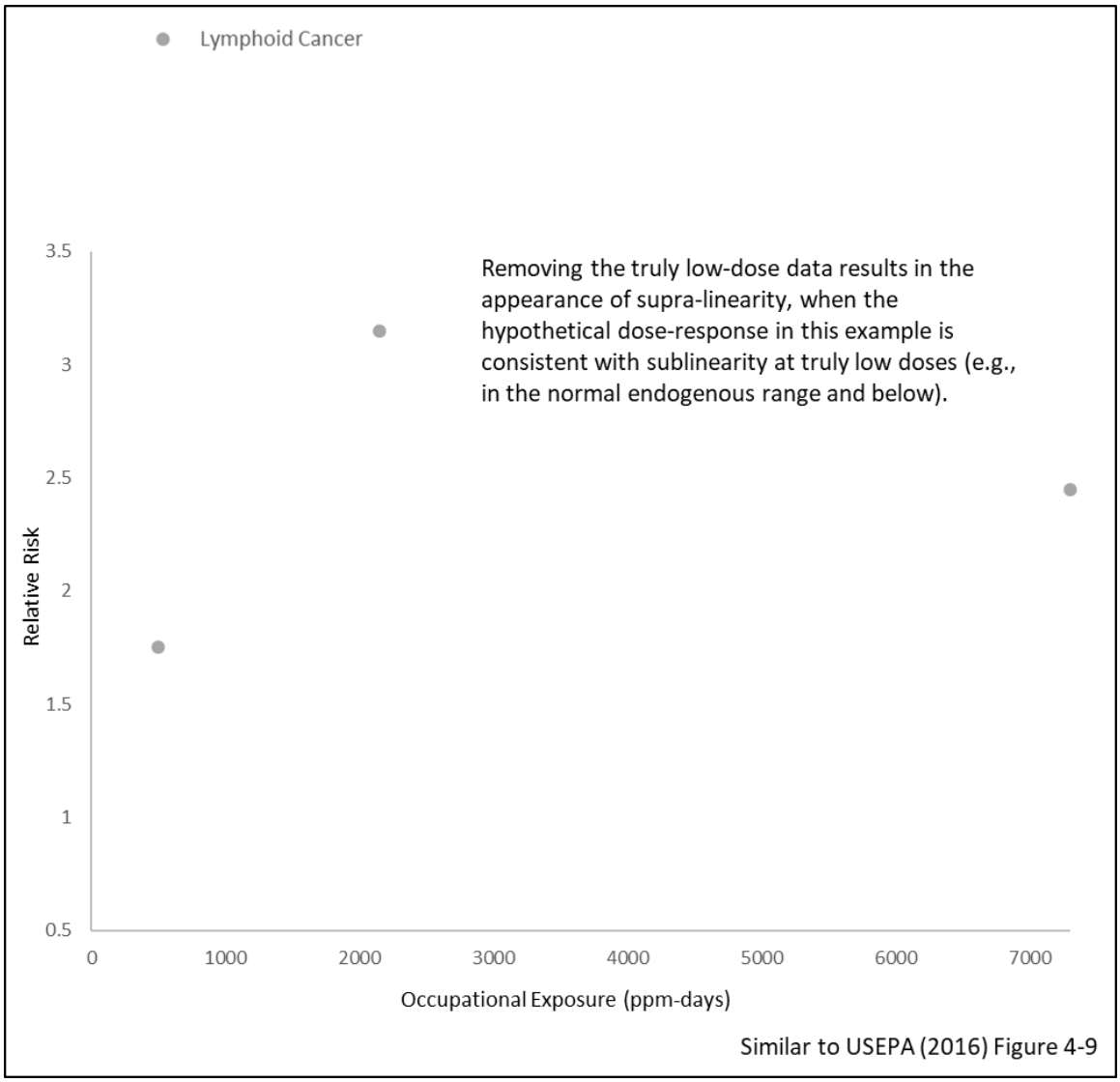


Figure 17: Seemingly Supra-linear Dose-Response from Removal of Hypothetical Low-Dose Data for Lymphoid Cancer

Figure 18 also depicts the possibility of a downward shift in the apparent dose-response curve in the absence of truly low-dose data, where the dose range for the apparent supra-linear curve on the left could be similar to that in Figure 14.

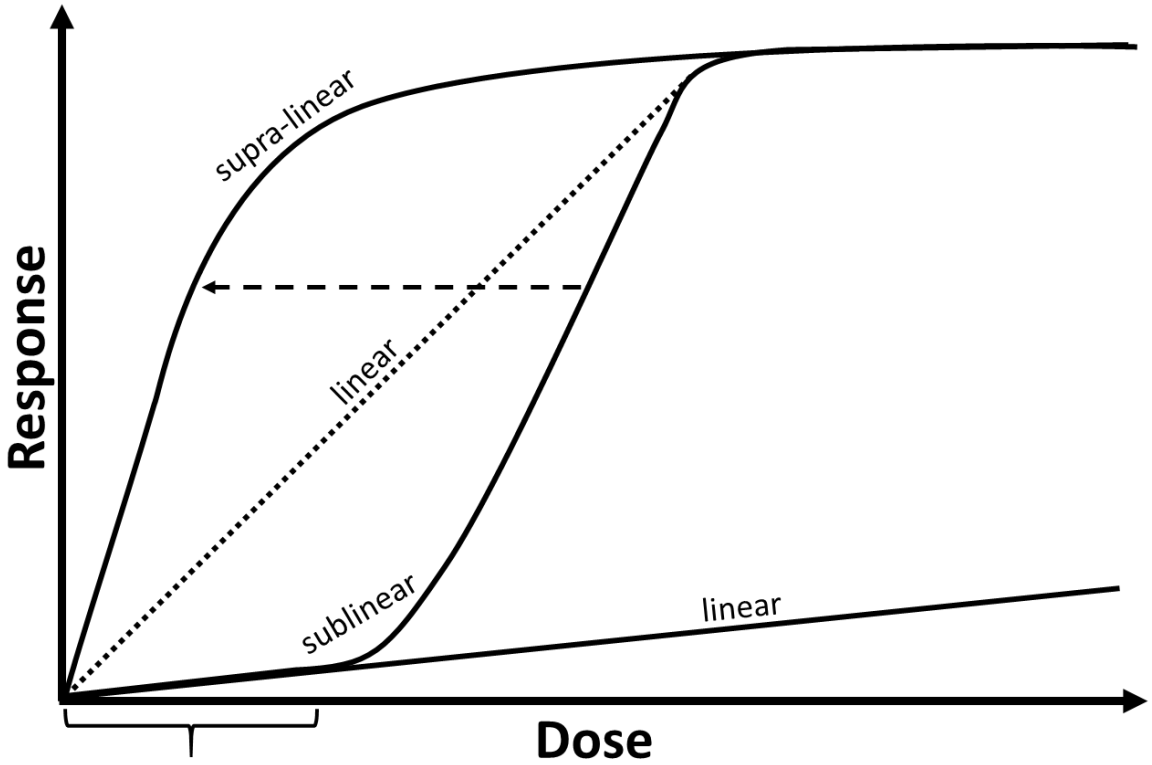


Figure 18: Seemingly Supra-linear Dose-Response from Removal of Hypothetical Low-Dose Data

These examples simply demonstrate the hypothetical possibility of the appearance of an overall supra-linear dose-response, despite an underlying true dose-response that is sublinear at truly low doses, *when available data are at relatively high doses* above the sublinear portion of the curve and into the steep slope portion where a high response per unit dose is induced.

To help put the high occupational EtO exposures into perspective, the TCEQ notes that the NIOSH cohort worker exposure means of 3.5-4.6 ppm (Hornung et al. 1994) are 778-1,022 times the air concentration corresponding to the 95th percentile of the normal endogenous background range (4.5 ppb; Table 4 of Kirman and Hays 2017) and over 1,000,000 times higher than central tendency environmental levels (i.e., background and environmental exposure means ≈0.0044-0.0062 µg/m^3 (0.0024-0.0034 ppb) per USEPA 2016). Even today the OSHA PEL (1 ppm) is 222 times the air concentration corresponding to the 95th percentile of the normal endogenous background range (4.5 ppb; Table 4 of Kirman and Hays 2017) and around

294,000-417,000 times higher than central tendency environmental levels (i.e., background and environmental exposure means ≈0.0044-0.0062 µg/m^3 (0.0024-0.0034 ppb) per USEPA 2016).

The TCEQ has not evaluated the hypothetical above further as it is somewhat beyond the scope of this DSD.

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

A4.1 Lymphoid Cancer:

A4.1.1 Corrected p-value example for the log-linear spline model with knot at 1,600 ppm-days

The likelihood ratio test is used to test whether a fitted model significantly improves the fit of the data by estimating parameters instead of just assuming a baseline (null) model for the data. The likelihood ratio test is evaluated by comparing the likelihood of the model with the estimated parameters and the likelihood of the null model. If the likelihood of the model with the estimated parameters is equal to the likelihood of the null model, then the natural logarithm of the ratio of these likelihoods multiplied by two follow a Chi-Square distribution with as many degrees of freedom as the number of parameters estimated for the fitted model. Thus, if the fit of the baseline (null) model and the model with estimated parameters are not different,

$$Chi - Square(k) = \chi_k^2 = -2\ln\left(\frac{likelihood\ for\ null\ model}{likelihood\ for\ fitted\ model}\right)$$

This can also be written as follows,

$$\chi_k^2 = -2LogL(null\ model) + 2LogL(fitted\ model)$$

Here *k* is the number of degrees of freedom (*k* is the number of parameters that were estimated in excess of the parameters estimated for the null model or nested model).

For the log-linear spline model with knot at 1,600 ppm-days for lymphoid cancer (Table D-33 on p. D-46 of USEPA 2016), the χ_k^2 value was equal to 5.2722 (463.912-458.640) and *k* was set to 2. This resulted in a p-value of 0.0716. That is, the fitted model was assumed to have two parameters; namely, the slope below the knot and the slope above the knot. The results are from a Statistical Analysis System (SAS) output for the model specified. *The two-piece log-linear model specified included a knot. This knot was determined so that the likelihood of the spline model was maximized.* That is, the knot is another parameter that was searched for outside SAS. *Because the estimation of the knot was done outside SAS, the SAS program did not count the knot as a parameter and, consequently, the Chi-Square test SAS reported does not reflect the fact that the knot was also estimated.* The correct Chi-Square that accounts for the fact that the knot was estimated outside SAS should then be 5.2722, but *k* (the degrees of freedom) should be three. This corrected calculation would result in a p-value of 0.1529. That is, the corrected p-value indicates that the likelihood of the log-linear spline model with knot at 1,600

ppm × days is not different from the likelihood of the null model at the 5% significance level. In plain words, there is not enough evidence indicating that the fitted two-piece log-linear spline model explains the variability in the data any better than the null model. *The same is true for the linear two-piece spline model with a "knot" at 1,600 ppm × days selected by USEPA (2016), which has a correct p-value of 0.14.*

A4.1.2 Corrected AIC value example for the log-linear spline model with knot at 1,600 ppm-days

The AIC is equal to 2*k* - 2*LogL* where *k* is the number of parameters estimated for the model and LogL is the logarithm of the likelihood. Table D-33 in USEPA (2016) lists the -2LogL as 458.640 and the AIC as 462.640. That is:

$$462.640 = 2k + 458.640$$

The AIC and –2LogL implies that k equals 2. That is, the spline model was assumed to have estimated two parameters; namely, the slope below the knot and the slope above the knot. The results in Table D-33 (p. D-46 of USEPA 2016) consist of SAS output for the two-piece log-linear spline model specified. The model specified included a knot. This knot was previously estimated using a separate optimization procedure outside the SAS run, so the likelihood of the model was maximized only *conditional on* the estimated knot-value used for that calculation. *Consequently, the knot must be treated as an additional parameter that was estimated outside SAS. However, because the estimation of the knot was done outside SAS, the SAS run performed by USEPA (2016) did not count the knot as a model parameter and, consequently, the resulting AIC value it obtained does not reflect that the knot was in fact estimated.* USEPA could have requested SAS to account properly for the extra degree of freedom properly associated with its estimated knot value, but USEPA evidently elected not to make this request of SAS.

The *correct AIC*, which accounts for the fact that the knot was estimated outside SAS, should instead be:

$$AIC = 464.640 = 2 \times 3 + 458.640$$

Correct AIC values and p-values for all models in Table 4-6 of USEPA (2016) are summarized in the corrected USEPA Table 4-6 below, which is Table 38 of this DSD (i.e., the p-values and AIC values have been corrected to reflect the degree of freedom for the knot in the two-piece spline models and to reflect the likelihood difference between SAS procedures used for linear and log-linear models).

Table 38: Corrected USEPA Table 4-6 - Models Considered for Modeling the EtO Exposure-Response Data for Lymphoid Cancer Mortality in Both Sexes in the NIOSH Cohort for the Derivation of Unit Risk Estimates

Model [a]	p-value [b]	AIC [c]	USEPA Comments
Two-piece spline models			
Linear spline model with knot at 1,600 ppm × days	0.14	*464.5*	**SELECTED.** Adequate statistical and visual fit, including local fit to low-exposure range; linear model; AIC within two units of lowest AIC of models considered.
Linear spline model with knot at 100 ppm × days	0.11	*463.8*	Good overall statistical fit and lowest AIC of two-piece spline models, but poor local fit to the low-exposure region, with no cases below the knot.
Log-linear spline model with knot at 1,600 ppm × days	0.15	464.6	Linear model preferred to log-linear (see text above).
Log-linear spline model with knot at 100 ppm × days	0.11	463.8	Good overall statistical fit and tied for lowest AIC of two-piece spline models, but poor local fit to the low-exposure region, with no cases below the knot.
Linear (ERR) models (RR = 1 + β × exposure)			
Linear model	0.13	*463.6*	Not statistically significant overall fit and poor visual fit.
Linear model with log cumulative exposure	0.02	*460.6*	Good overall statistical fit, but poor local fit to the low-exposure region.
Linear model with square-root transformation of cumulative exposure	0.053	*462.2*	Borderline statistical fit, but poor local fit to the low-exposure region.
Log-linear (Cox regression) models (RR = $e^{\beta \times exposure}$)			
Log-linear model (standard Cox regression model)	0.22	464.4	Not statistically significant overall fit and poor visual fit.
Log-linear model with log cumulative exposure	0.02	460.4	Good overall statistical fit; lowest AIC[c] of models considered; low-exposure slope becomes increasingly steep as exposures decrease, and large unit risk estimates can result; preference given to the two-piece spline models because they have a better ability to provide a good local fit to the low-exposure range.
Log-linear model with square-root transformation of cumulative exposure	0.08	462.8	Not statistically significant overall fit and poor visual fit.

[a] All with cumulative exposure as the exposure variable, except where noted, and with a 15-yr lag.
[b] p-values from likelihood ratio test, except for linear regression of categorical results, where Wald p-values are reported. $p < 0.05$ considered "good" statistical fit; $0.05 < p < 0.10$ considered "adequate" statistical fit if significant exposure-response relationships have already been established with similar models.
[c] AICs for linear models are directly comparable and AICs for log-linear models are directly comparable. However, for the lymphoid cancer data, SAS proc NLP (where NLP = nonlinear programming) consistently yielded −2LLs and AICs about 0.4 units lower than proc PHREG for the same models, including the null model, presumably for computational processing reasons, and proc NLP was used for the linear RR models. Thus, AICs for linear models

are equivalent to AICs about 0.4 units higher for log-linear models. No AIC was calculated for the linear regression of categorical results. *In order to make the AICs comparable for different models, the AICs for the linear models have been increased by 0.4 to reflect the discrepancy in the -2LogL values reported by the SAS proc NLP and by SAS PHREG (as italicized in this table).*

Table 38 shows that *neither the linear two-piece spline model with a "knot" at 1,600 ppm × days selected by USEPA (2016) nor the standard Cox regression model fit the data statistically significantly better than the null model* (zero slope). Additionally, the AIC values are very similar. *However, as use of an overall supra-linear model (i.e., the steep lower-dose slope) is not scientifically justified (see Section 3.4.1.4.1), the two-piece spline models are not considered for adoption*; nor are other models that have an inherently supra-linear dose-response over the exposure range (i.e., log-linear or linear models with log cumulative exposure or with square-root transformation of cumulative exposure). As for the linear model, it neither fits the data statistically better than the null model (at the 5% significance level) nor is consistent with USEPA's/TCEQ's expectation of sublinearity in the endogenous range, while the standard Cox regression model is consistent. Lastly, *no superior model fit is readily apparent visually based on accurate depictions of model fit to the actual underlying data* (Appendix 5). For reasons discussed in Section 3.4.1.4.2, *the TCEQ selects the standard Cox regression model for lymphoid cancer mortality*.

A4.2 Breast Cancer Incidence

A4.2.1 Corrected AIC example for the linear spline model with knot at 5,750 ppm-days

Similar to Table 38 above for lymphoid cancer, correct AIC values and p-values for all breast cancer incidence models in Table 4-14 of USEPA (2016) are summarized in the corrected USEPA Table 4-14 below, which is Table 39 in this DSD (i.e., the p-values and AIC values have been corrected to reflect the degree of freedom for the knot in the two-piece spline models and to reflect the likelihood difference between SAS procedures used for linear and log-linear models).

Table 39: Corrected USEPA Table 4-14 - Models Considered for Modeling the EtO Exposure-Response Data for Breast Cancer Incidence in Females in the Subcohort with Interviews from the NIOSH and Health Incidence Study Cohort for the Derivation of Unit Risk Estimates

Model [a]	p-value [d]	AIC [b]	USEPA Comments
Two-piece spline models			
Two-piece linear spline model (knot at 5,750 ppm × days)	*0.0367*	1,956.360 [e]	**SELECTED**. Good overall statistical fit and good visual fit, including local fit to low-exposure range; linear model; AIC within two units of lowest AIC of models considered.
Two-piece log-linear spline model (knot at 5,800 ppm × days)	*0.0384*	*1,956.485*	Good overall statistical fit and good visual fit, including local fit to low-exposure range; preference given to the two-piece linear spline model primarily because it has the advantageous property of linearity, but it also has a marginally better statistical fit (lower AIC).
Linear (ERR) models (RR = 1 + β × exposure)			
Linear model with square-root transformation of cumulative exposure	*0.0038*	1,952.501	Good overall statistical fit and lowest AIC; low-exposure slope becomes increasingly steep as exposures decrease, and large unit risk estimates can result; preference given to the two-piece spline models because they have a better ability to provide a good local fit to the low-exposure range.
Linear model with untransformed cumulative exposure	*0.0114*	1,954.526	Good overall statistical fit but poorer local fit to low-exposure range than the two-piece spline models; higher AIC than selected model.
Log-linear (Cox regression) models ($RR = e^{\beta \times exposure}$)			
Log-linear model with square-root transformation of exposure	*0.0049*	1,953.028	Good overall statistical fit; low-exposure slope becomes increasingly steep as exposures decrease, and large unit risk estimates can result; preference given to the two-piece spline models because they have a better ability to provide a good local fit to the low-exposure range.
Log-linear model with (natural) log cumulative exposure	*0.0302*	1,956.176	Good overall statistical fit but poor local fit to low-exposure range; low-exposure slope becomes increasingly steep as exposures decrease, and large unit risk estimates can result; higher AIC than selected model.
Log-linear model (standard Cox regression)	0.0404	1,956.675	Good overall statistical fit but poor local fit to low-exposure range (too shallow); AIC exceeds that of selected model by >2.
Linear regression of categorical results			
Linear regression of categorical results, excluding the highest exposure quintile	---	--- [c]	Not statistically significant, as one might expect because the approach, which is based on categorical data, has low statistical power; preference given to models that treated exposure as a continuous variable and that also provided reasonable representations of the low-exposure region.

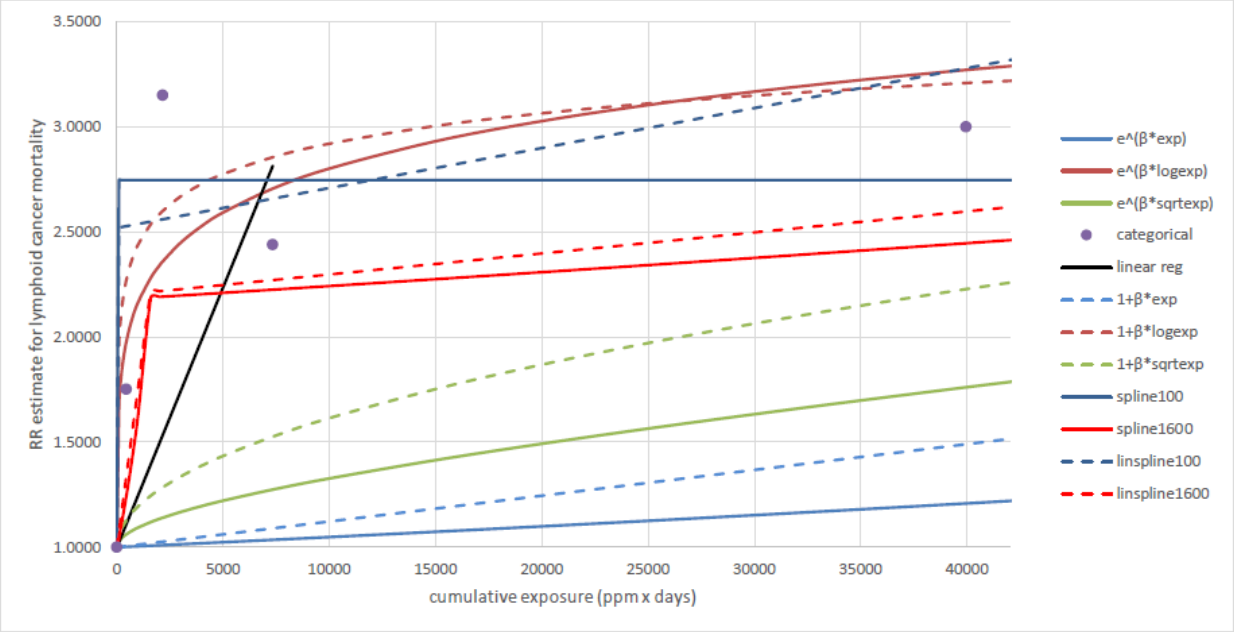
[a] All with cumulative exposure as the exposure variable, except where noted, and with a 15-yr lag, and all with exposure as a continuous variable except for the linear regression of categorical results.
[b] AIC = 2p-2LL, where *p* = number of parameters and LL = ln(likelihood), assuming two exposure parameters for the two-piece spline models.
[c] Not calculated.

[d] *p-values were calculated from EPA's Table D-2.*
[e] *AIC values for the two-piece spline models were adjusted to reflect the degree of freedom for the knot.*

Table 39 shows that *both the linear two-piece spline model with a "knot" at 5,750 ppm × days selected by USEPA (2016) and the standard Cox regression model selected by the TCEQ fit the data statistically significantly better than the null model* (zero slope). Additionally, the AIC values are very similar. *However, as use of an overall supra-linear model (i.e., the steep lower-dose component) is not scientifically justified (see Section 3.4.1.4.1), the two-piece spline models are not considered for adoption*; nor are other models that have an inherently supra-linear dose-response over the exposure range (i.e., log-linear or linear models with log cumulative exposure or with square-root transformation of cumulative exposure). *While the TCEQ considered standard Cox regression modeling results for breast cancer incidence, for reasons discussed in Section 3.4.1.6.1 and Appendix 6 (e.g., the epidemiological weight of evidence and inability of animal data to support breast cancer as a cancer site of in humans) the agency selected lymphoid cancer as the critical cancer endpoint*.

Appendix 5 Visual Fit to the Underlying NIOSH Data

Visual fit to the data was used by USEPA (2016) as a criterion for model selection. However, no appropriate visual comparison of model fit to the lymphoid cancer mortality data can be made based on Figure 4-3 (p. 4-21 of USEPA 2016) since *the data shown are not even the data to which the models were fit*. As such, USEPA Figure 4-3 (shown below as Figure 19 of this DSD) misrepresents model fit.



e^(β*exp): $RR = e^{(\beta \times exposure)}$; e^(β*logexp): $RR = e^{(\beta \times \ln(exposure))}$; e^(β*sqrtexp): $RR = e^{(\beta \times sqrt(exposure))}$; categorical: $RR = e^{(\beta \times exposure)}$ with categorical exposures, plotted at the mean cumulative exposure; linear reg: weighted linear regression of categorical results, excluding highest exposure group (see text); 1 + β*exp: $RR = 1 + \beta \times exposure$; 1 + β*logexp: $RR = 1 + \beta \times \ln(exposure)$; 1 + β*sqrtexp: $RR = 1 + \beta \times sqrt(exposure)$; spline100(1,600): Two-piece log-linear spline model with knot at 100 (1,600) ppm × days (see text); linspline100(1,600): Two-piece linear spline model with knot at 100 (1,600) ppm × days (see text). (Note that, with the exception of the categorical results and the linear regression of the categorical results, the different 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. They are, however, comparable in terms of general shape.)

Source: Steenland reanalyses for males and females combined; see Appendix D (except for linear regression of categorical results, which was done by EPA).

Figure 4-3. Exposure-response models for lymphoid cancer mortality vs. occupational cumulative exposure (with 15-year lag).

Figure 19: USEPA (2016) Figure 4-3

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

Figure 19 reproduces Figure 4-3 in USEPA's 2016 risk assessment. This figure shows the rate ratios of twelve models. Eleven of those models have a parametric functional form and one model (labeled here "categorical") estimates non-parametric rate ratios of the lymphoid mortality grouped by quintiles. Each quintile summarizes information for 11 lymphoid deaths (9

in the non-exposed quintile). *The "categorical" points are not the data* – they are estimates of the rate ratios. *Rate ratios are not observed, they are estimated*. Furthermore, the non-parametric rate ratios derived by USEPA and shown in Figure 19 do not show the full range of possible rate ratios and cumulative exposures. Table D-28 of USEPA (2016) includes the uncertainty (i.e., 95% CIs) around USEPA's "categorical" RRs and is reproduced here as Table 40 for lymphoid cancer (males and females combined).

Table 40: Lymphoid Cancer Categorical RRs and 95% CIs (male + female)

Cumulative exposure range, 15-year lag (ppm-days)	Mean* Cumulative Exposure (ppm-days)	Rate Ratio	Lower Confidence Limit on the Rate Ratio	Upper Confidence Limit on the Rate Ratio
0 (lagged out)	0	1.00	--	--
>0 – 1,200	446	1.75	0.59	5.25
1,201 – 3,680	2,143	3.15	1.04	9.49
3,681 – 13,500	7,335	2.44	0.80	7.50
>13,500	39,927	3.00	1.02	8.45

Categorical rate ratios (RRs) should not be used for visually comparing models fit to individual data, particularly when appropriate statistical model fit criteria are available. More specifically, estimated nonparametric RRs are calculated with respect to an underlying background hazard rate that is also estimated nonparametrically. The RRs of parametric models fit to the individual data are defined with respect to an underlying background hazard rate estimated by the model. However, the underlying background hazard rates estimated by the nonparametric RRs and the parametric model are generally different. *A better comparison of models fit to the observed data is to use the predictiveness of the model; that is, the capability of the model to estimate the observed number of deaths with a certain degree of confidence (see Appendix 2).* Moreover, visual interpretation of the consistency of categorical RRs with the shape/slope of a modelled dose-response can change as the number of exposure categories changes. For example, Figures 1-3 of Valdez-Flores and Sielken (2013) demonstrate, among other things, how the dose-response (i.e., dose-RR) slope for breast cancer mortality in the NIOSH cohort appears very steep when compared to only four exposure categories but seems more shallow when additional categories are added (i.e., up to 20 and 61 categorical RRs). In the present case, the overall dose-response appears ill represented by only a few categorical RRs, whether for breast cancer (see Figures 1-3 of Valdez-Flores and Sielken 2013) or lymphoid cancer (see below and supplementary material for Valdez-Flores and Sielken 2013).

The visual presentation of only a few exposure categories can blind the data user to the variability in the underlying dose-response data, and by corollary, preclude an appropriate

visual assessment/comparison of model fit to the actual individual data. Figure 20 below shows the same models in Figure 19 with the superposition of the estimated RRs (open circles labeled as categorical in Figure 20 and USEPA's nonparametric estimates labeled as USEPA's 5 RRs shown as red dots). Figure 20 shows the vertical axis using a multiplicative scale to show resolution and show the full range of rate ratios while the x-axis shows the full range of exposure of lymphoid decedents in the NIOSH study. Figure 21 is the same as Figure 20 but restricted to the rate ratios (vertical scale) and to the cumulative exposures (x-axis) used by USEPA (2016) in their Figure 4-3. Figures 20 and 21 also show a dotted line that fits an exponential model to the individual (categorical RRs shown as open circles). The intercept of this line can be used to approximate the ratio of the underlying background hazard rate implied by the standard Cox proportional hazards model to the underlying background hazard rate implied by the nonparametric estimates. Finally, Figure 21 shows USEPA's selected model and the standard Cox proportional hazards model after adjusting the intercept for the differences in the estimated baseline risks. The standard Cox proportional hazards model in Figure 22 (dashed blue line) is no longer a RR function, but rather adjusts for discrepancies in the estimated baseline risks of two models so that they can be visually compared in the same graph.

In looking at all lymphoid cancer death RRs for the NIOSH cohort in Figure 20, Figure 21, and Figure 22 below (e.g., as opposed to a few categorical RRs represented by the red dots), *objective examination of the model fits to the underlying data reveals no readily apparent superior fit by any particular model*. What is most readily apparent is the loss of visualized information that results from only using the five grouped RRs (represented by the red dots) as in Figure 4-3 of USEPA (2016). The nonparametric rate ratios for individual cases (categorical) represented by the black circles in Figure 22 below form no discernable pattern that appears most consistent with any specific model (i.e., visual fit cannot be used to readily identify a model fit most representative of the actual data). In fact, other dose-responses could be added that would appear equally plausible and/or consistent with these high-dose occupational data.

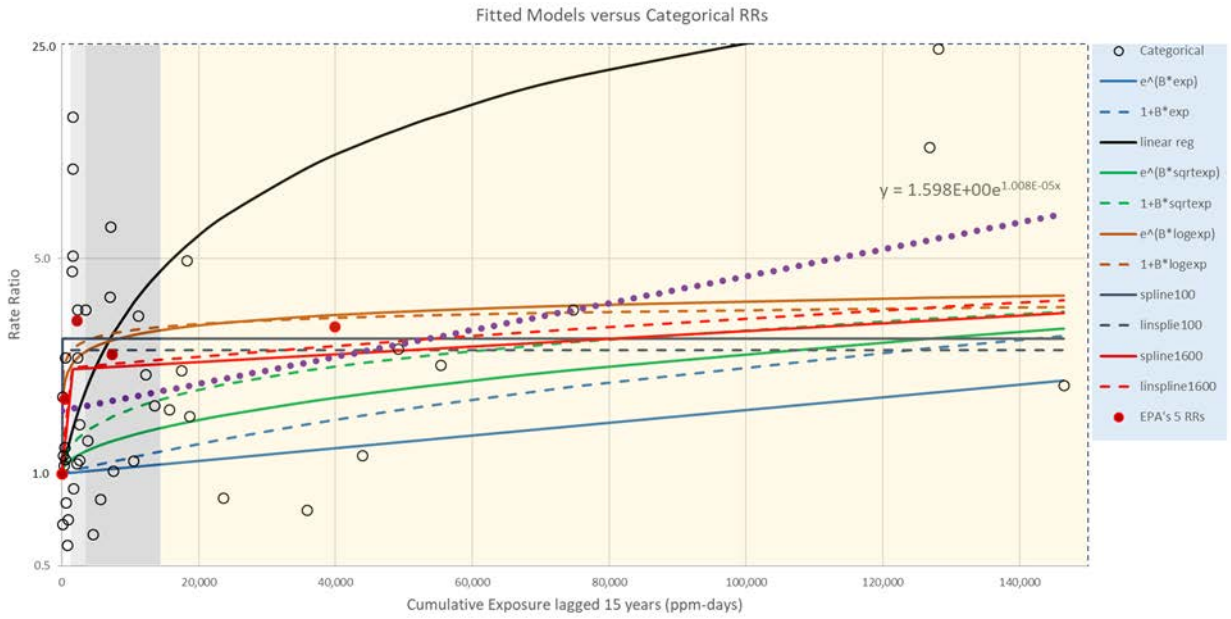


Figure 20: Lymphoid Cancer Death Categorical Rate Ratios (RRs) and Various Fitted Models for 15-Year Lagged Occupational Doses ≤150,000 ppm × days (NIOSH cohort)

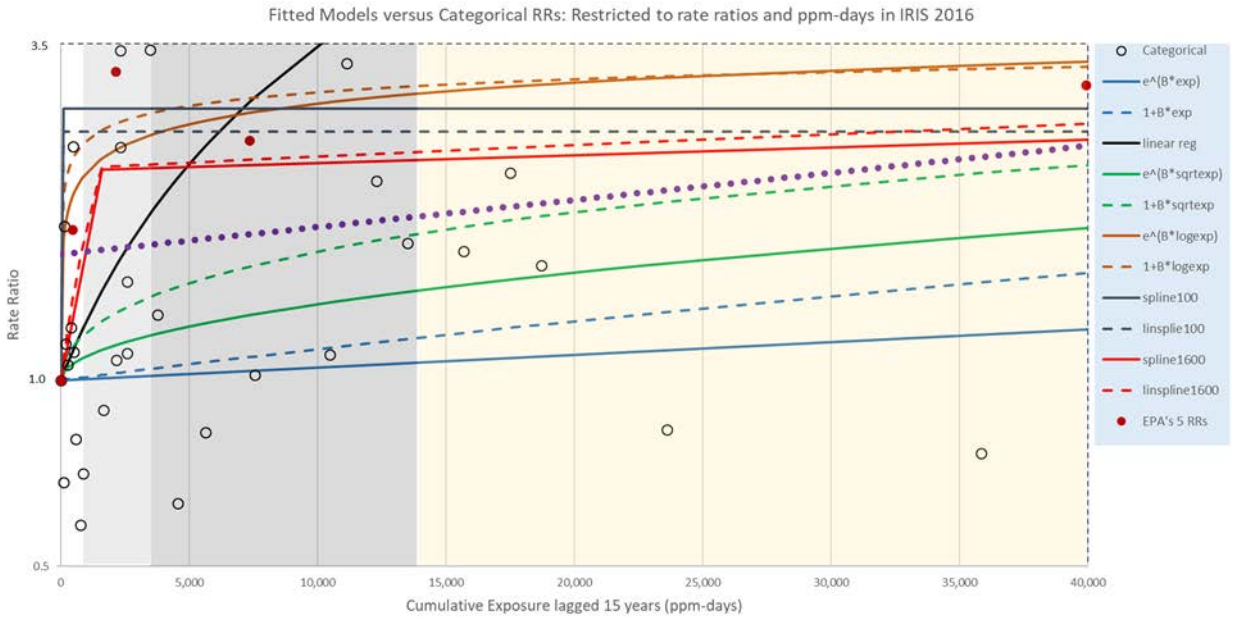


Figure 21: Lymphoid Cancer Death Categorical RRs and Various Fitted Models for 15-Year Lagged Occupational Doses ≤40,000 ppm × days (NIOSH cohort)

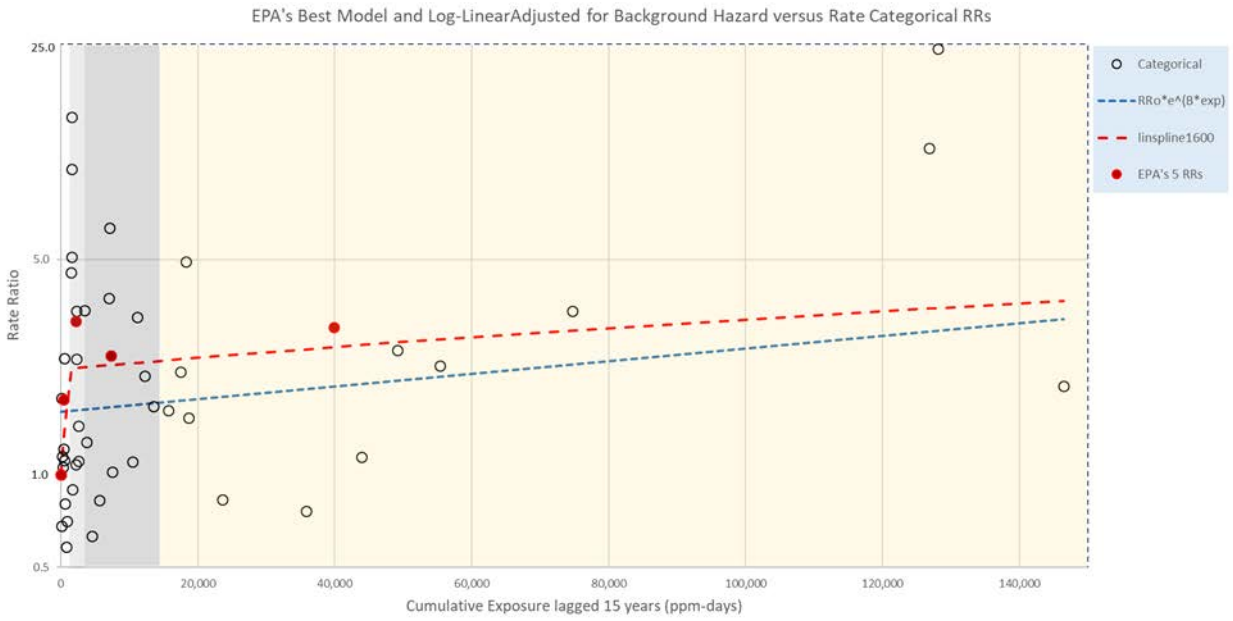


Figure 22: Lymphoid Cancer Death Categorical RRs and the Cox Proportional Hazards and Two-Piece Spline ("knot" at 1,600 ppm × days) Fitted Models for 15-Year Lagged Occupational Doses ≤150,000 ppm × days (NIOSH cohort)

[Note: In Figure 22, the dotted light blue line approximates the correct visual representation of the log-linear model (standard proportional hazards model) fit to the full NIOSH dataset after adjusting for the difference in baseline risks between the rate ratios and the log-linear model, thereby addressing USEPA's following footnote to Figure 4-3 (p. 4-21 of USEPA 2016) concerning the visual incomparability of model fit to the data, "Note that, with the exception of the categorical results and the linear regression of the categorical results, the different 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." The model "RRo* e^(B*exp)" is an approximation of the log-linear model (e^(B*exp)) adjusted through multiplying by the ratio of the underlying baseline hazard rate of the model to the underlying baseline hazard rate the nonparametric estimates.]

Because the plotting of TCEQ's log-linear model in Figure 22 accounts for the different implicit estimated baseline risks of the non-parametric RRs and the log-linear model on the y-axis, the figure *shows a more fair visual comparison of the fits* of the TCEQ log-linear model, the USEPA two-piece linear spline model (USEPA's selected model for risk evaluation), and the individual estimated RRs. In doing so, *Figure 22 shows that the RRs estimated by the log-linear model (TCEQ's selected model) are certainly as visually consistent with the RRs estimated non-parametrically (i.e., the open circles in Figure 20, Figure 21, and Figure 22) as USEPA's model*.

Misinterpretation in the comparison of parametric and categorical (non-parametric) RRs used to judge model fit has been published in the peer-review literature (e.g., Valdez-Flores and Sielken 2013).

These more elucidating, transparent, and accurate visual depictions of model fit reveal no readily apparent superior model fit (e.g., see the three figures above and Figure 3 of Valdez-Flores and Sielken 2013). Regardless, the TCEQ uses a better approach to judge model fit to the observed data. Model fit is best judged by appropriate statistical model fit criteria and the ability to predict the underlying data modeled, which are evaluated elsewhere in this DSD (Sections 3.4.1.2.2.3, 3.4.1.3, Appendix 2 and Appendix 4). **Briefly, correctly calculated p-values and AIC values indicate that the standard Cox proportional hazards model (TCEQ's selected model) fits the data as well as USEPA's unconventional two-piece spline model and is reasonably accurate in predicting the underlying lymphoid cancer data (i.e., nether significantly over- or under-predicting) whereas USEPA's selected model assessment statistically significantly over-predicts the number of lymphoid cancer mortalities both on a whole cohort and exposure quintile-specific basis.**

In regard to the alleged sharp rise in excess risk that appears when using five categorical RRs as in Figure 4-3 of USEPA (2016) and Figure 19 through Figure 22 above (represented by red dots): (1) visual representation of summary statistics can be misleading when the summary statistics are believed to be observations; and (2) summarizing the RRs by using fewer grouped individual cases only masks the true variability in the underlying estimates of the RRs. Table 4-2 in the USEPA (2016) risk assessment lists the estimates of the RRs (ratios of the hazard rate for each exposure quintiles to the hazard rate for the unexposed workers). The RRs are summary estimates of the estimated individual rate ratios shown by circles in Figure 20, Figure 21, and Figure 22 and are approximately located in the center of the 11 individual rate ratios included in each quintile. Table 41 below shows USEPA's quintile RRs (USEPA calls them ORs) with their corresponding 95% CIs along with the average RR of the 11 individual RRs and the range of the individual RRs.

Table 41: USEPA Quintile RRs and 95% Confidence Intervals versus Corresponding Quintile-Specific Individual RRs

Quintile	USEPA's Quintile RRs [1] (95% Confidence Interval)	Average of 11 [2] Individual RRs in the Quintile	Individual RRs in the Quintile [3]
2	1.75 (0.59, 5.25)	1.46	0.58, 0.68, 0.71, 0.80, 1.06, 1.11, 1.15, 1.22, 1.77, 2.38, 4.55
3	3.15 (1.04, 9.49)	4.04	0.89, 1.08, 1.11, 1.28, 1.44, 2.38, 3.41, 3.42, 5.11, 9.82, 14.49

Quintile	USEPA's Quintile RRs [1] (95% Confidence Interval)	Average of 11 [2] Individual RRs in the Quintile	Individual RRs in the Quintile [3]
4	2.44 (0.80, 7.50)	2.22	0.63, 0.82, 1.02, 1.10, 1.62, 1.67, 2.10, 2.16, 3.25, 3.75, 6.34
5	3.00 (1.02, 8.45)	4.99	0.76, 0.83, 1.14, 1.53, 1.94, 2.26, 2.54, 3.40, 4.93, 11.50, 24.11

[1] Source: Table 4-2 of USEPA's (2016) risk assessment report.
[2] The average of the 11 individual RRs are not statistically significantly different than the quintile RRs estimated by USEPA.
[3] Most individual rate ratios are inside the 95% confidence interval of the RR corresponding to the quintile.

Figure 22 and this table show that the alleged steep increase at low cumulative exposures and plateauing of the RRs at higher cumulative exposures is an artifact of summarizing the RRs into quintiles. The 95% CIs of the quintile RRs and the individual RRs based on each lymphoid decedent shown in the table represent the variability in the NIOSH data for lymphoid cancer. *The alleged supra-linearity (steep increase for low cumulative exposures and plateauing at higher cumulative exposures concluded from the red dots in Figure 22) is not supported by the individual RRs (open circles) in Figure 22, which form no discernable dose-response pattern*. This figure shows that the two models fit the individual RRs about the same. This is *corroborated by the p-values and AICs in Table 38* where the linear and the standard Cox proportional hazards model have preferable (i.e., lower) AIC values once the correct degrees of freedom (*df*) for USEPA's selected model are correctly accounted for. **It is important to recognize that standard statistical measures of model fit are calculated so that visual fit need not be relied upon, although visual examination of the actual individual data is consistent with the correctly calculated p-values and AIC values that indicate that contrary to USEPA (2016) assertions, their overall supra-linear two-piece spline model does not fit the data better than the TCEQ's standard Cox proportional hazards model. Moreover and again, TCEQ's selected model assessment reasonably accurately predicts the underlying lymphoid cancer data whereas USEPA's selected model assessment statistically significantly over-predicts the number of lymphoid cancer mortalities both on a whole cohort and exposure quintile-specific basis (see Section 3.4.1.2.2.3 and Appendix 2).**

A5.2 Model-Specific Implicitly Estimated Baseline Risks

USEPA's footnote to several figures indicates that the different models and the non-parametric RRs cannot be compared along the y-axis because "the different models have different implicitly estimated baseline risks." USEPA is correct. All models in Figure 4-3 of USEPA (2016) risk assessment (Figure 19 herein), with the exception of the "linear reg" model, are fit to hazard rates (not fit to RRs). The functional form of all models is

$$HR_i(d) = HR_i(0) \times f_i(d)$$

where $HR_i(d)$ is the hazard rate of model i at cumulative exposure *d,* $HR_i(0)$ is the "estimated baseline risk" for model i, and $f_i(d)$ is the function of the relative risk at cumulative exposure *d* for model i.

Note that by dividing $HR_i(d)$ by the "estimated baseline risk" $HR_i(0)$, the function $f_i(d)$ is the relative risk at cumulative exposure *d* for model i. Note also, that each model i could result in different estimates of the baseline risk, $HR_i(0)$. That means, all models would have relative risk ($f_i(0)$) equal to 1 at cumulative exposure equal to 0. However, the "estimated baseline risk" $HR_i(0)$, could be very different. The model for USEPA's 5 categorical RRs, USEPA's two-piece linear spline model, and TCEQ's standard Cox proportional hazards model have the following functional forms:

Model 1 ("EPA's 5 RRs" and "Individual RRs" in the figures): The non-parametric model fit to the data is given by the expression

$$HR_{NP,k}(d) = HR_{NP}(0) \times RR_{NP,k}(d)$$

where $HR_{NP,k}(d)$ is the hazard rate for the *k*-th group at mean cumulative exposure *d*, $HR_{NP}(0)$ the "estimated baseline risk" for the nonparametric model, and $RR_{NP,k}(d)$ the relative risk for the *k*-th group. Although the function does not depend on the magnitude of the exposure *d,* the function is written with the *d* for the sake of consistency. (USEPA expresses the function $RR_{NP,k}(d) = e^{\beta_k \times d}$ where "*d*" is a "categorical exposure." Using USEPA's expression guarantees $RR_{NP,k}(d)$ is non-negative when doing a search.)

Model 2 ("linspline1600" in the figures): The functional form of USEPA's selected model (linspline1600) two-piece linear model is

$$HR_{spl}(d) = HR_{spl}(0) \times \begin{cases} 1 + \beta_1 \times d & d \leq knot \\ 1 + \beta_1 \times d + \beta_2 \times (d - knot) & d > knot \end{cases}$$

where $HR_{spl}(d)$ the hazard rate at cumulative exposure *d*, $HR_{spl}(0)$ the "estimated baseline risk" for the two-piece linear model, $1 + \beta_1 \times d$ is the relative risk at cumulative exposures *d* below the *knot*, $1 + \beta_1 \times d + \beta_2 \times (d - knot)$ is the relative risk at cumulative exposures *d* above the *knot*, and *knot* is the cumulative exposure where the slope of the relative risk changes. USEPA estimated the *knot* at 1,600 ppm-days.

Model 3 ("e^(β*exp)" in the figures): The functional form of TCEQ's selected model ($e^{\beta * exp}$) standard Cox proportional hazards model is

$$HR_{Cox}(d) = HR_{Cox}(0) \times e^{\beta \times d}$$

where $HR_{Cox}(d)$ the hazard rate at cumulative exposure *d*, $HR_{Cox}(0)$ the "estimated baseline risk" for the standard Cox proportional hazards model, $e^{\beta \times d}$ is the relative risk at cumulative exposure *d*.

The relative risks (RRs) from each of the models described above are, by definition, equal to one at zero cumulative exposures. However, as indicated by USEPA's 2016 assessment and shown above for Models 1, 2, and 3, the "implicitly estimated baseline risks" ($HR_{NP}(0)$, $HR_{spl}(0)$, and $HR_{Cox}(0)$, for Models 1, 2, and 3, respectively) are different. That is, the RRs for the models cannot be compared for non-zero cumulative exposures without accounting for the differences in the "implicitly estimated baseline risks" ($HR_{NP}(0)$, $HR_{spl}(0)$, and $HR_{Cox}(0)$). The partial likelihood methodology for the proportional hazards models described above do not explicitly estimate the baseline risks ($HR_{NP}(0)$, $HR_{spl}(0)$, and $HR_{Cox}(0)$) and they are unknown. However, an approximation of the ratio of the "implicitly estimated baseline risks" for Models 2 and 3 to the "implicitly estimated baseline risks" for Model 1 ($HR_{spl}(0)/HR_{NP}(0)$ and $HR_{Cox}(0)/HR_{NP}(0)$, respectively) can be estimated from the non-parametric RRs based on the individual lymphoid decedents (open circles in Figures 20).

A5.3 Adjusting Models for Differences in Implicitly Estimated Baseline Risks for More Appropriate Visual Comparison

The ratio $HR_{spl}(0)/\ HR_{NP}(0)$ for Model 2 and $HR_{Cox}(0)/\ HR_{NP}(0)$ for Model 3 were calculated using weighted least squares and the corresponding RR functions for models 2 and 3, respectively. The best intercepts (ratios of baseline risk for each of the models to the baseline risk implied by the non-parametric RR estimates) multiply the rate ratio functions for Models 2 and 3. These adjusted Models 2 and 3 account for the differences in the baseline risks implied by the models and the implicitly estimated non-parametric baseline risks.

Figure 22 adjusts the standard Cox model (e^(β*exp)) by the estimated ratio $HR_{Cox}(0)/\ HR_{NP}(0)$. This adjusted plot is more appropriate to compare.

The y-axis in Figure 22 has been re-labeled to indicate that the models are normalized to the baseline risk implied by the non-parametric model rather than the models' own implied baseline risks. Figure 22 is divided into four regions using different colors. Each color shows the range of "individual RRs" and range of cumulative exposures that are summarized in each of "EPA's 5 RRs."

[That is, the RR for the highest quintile of "EPA's 5 RR" (red dots) is equal to 3 and is placed at a cumulative exposure of 39,927 ppm-days. The table above and Figure 22 show that the RR for the fifth quintile summarizes the individual RRs for the 11 lymphoid cancer decedents (open

circles) that had cumulative exposures greater than 13,500 ppm-days. Similarly, the RR for the fourth quintile summarizes the 11 individual RRs (open circles) based on lymphoid decedents with cumulative exposure between 3,681 and 13,500 ppm-days. The RR for the third quintile summarizes the 11 individual RRs (open circles) based on lymphoid decedents with cumulative exposure between 1,201 and 3,680 ppm-days. Finally, the RR for the second quintile summarizes the 11 individual RRs (open circles) based on lymphoid decedents with cumulative exposure greater than zero and less than or equal to 1,200 ppm-days.]

Figure 22 shows that the model selected by USEPA ("linspline1600") cannot be visually judged to be better than TCEQ's model ("e^(β*exp)").

In summary, although a secondary consideration to statistical analyses, visual comparisons of USEPA and TCEQ selected models fit the individual RRs approximately the same once differences in "baseline risks" of different RR models are reconciled. This conclusion is consistent with the conclusions drawn using correctly calculated standard model fit criteria (Appendix 4). If anything is to be gleaned from correctly calculated standard model fit criteria, Table 38 of the TCEQ DSD shows that after correcting for a missing degree of freedom, USEPA's selected model has an AIC of 464.5 and TCEQ's standard Cox proportional hazards model has an AIC of 464.4. This 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. Consistent with this, model performance in predicting the actual number of lymphoid cancers in the cohort as a whole and in each quintile demonstrates the superiority of the Cox proportional hazards model (Appendix 2).

Appendix 6 Weight of Evidence Regarding EtO-Induced Breast Cancer in Humans

Breast cancer requires a more detailed weight of evidence evaluation. USEPA (2016) acknowledges that the human data for EtO-induced breast cancer are less than convincing, which is remarkable given the extraordinarily high occupational exposure concentrations to a USEPA-purported extremely potent breast cancer carcinogen. This candidate cancer endpoint requires further evaluation given the particularly weak epidemiological evidence (i.e., USEPA-acknowledged inadequate human evidence) and laboratory animal data of questionable direct human relevance (see Sections 3.2.2 and 3.4.1.5.1).

A6.1 Epidemiological Data Relevant to EtO Exposure and Breast Cancer

The weight of evidence based on Table 42 below is that the SIRs/SMRs across individual EtO studies of breast cancer are *consistently not statistically significantly elevated*, most being less than 1. [Table 42 uses external referents for individual studies, as internal analyses appear not to be scientifically justified for breast cancer (next section).] Considering these results, it is not surprising that two recent meta-analyses of EtO studies that have examined breast cancer reported meta-RRs of 0.97 (0.80, 1.18) (Marsh et al. 2019) and 0.92 (0.84, 1.02) (Vincent et al. 2019). The Marsh et al. study concluded [*emphasis added*], "Evaluations of workers exposed during sterilization processes *do not support the conclusion that EO exposure is associated with an increased risk of breast cancer*." Similarly, the Vincent et al. (2019) study concluded, "Higher quality epidemiological studies demonstrated no increased risk of breast cancers."

Table 42: Human Studies Relevant to the Breast Cancer Weight of Evidence

Study Type	Workers (n)	EtO Exposure Level (ppm)	Observed (O)	Expected (E) [a]	O/E (95% CI)
Individual Studies					
Steenland et al. (2003)	7,576 female workers	Median ≈14 ppm-years; Mean >1 ppm [b]	230 [c]	258.4	**0.89** [d] (0.78, 1.01)
Steenland et al. (2004)	18,235 workers (≈55% female)	Mean of 26.9 ppm-years	103	102	**0.99** (0.84, 1.17)
	only female workers				**0.99** [e] (0.81, 1.20)
Mikoczy et al. (2011)	2,046 workers	Means ≤1.11 ppm;	33	38.54	**0.86** [f] (0.59, 1.20)

Study Type	Workers (n)	EtO Exposure Level (ppm)	Observed (O)	Expected (E) [a]	O/E (95% CI)
	(≈60% female) 615 female 287 female 295 female	Peaks up to 40-75 ppm Mean of 0.02 ppm in lowest cumulative exposure group Mean of 0.021 ppm in middle cumulative exposure group Mean of 1.11 ppm in highest cumulative exposure group			**0.52** [g] (0.25-0.96) 1.06 (**0.58**, 1.78) 1.12 (**0.65**, 1.79)
Norman et al. (1995)	928 female	TWA 50-200 ppm; 5-20 ppm post-corrective action 1980	12	7.64	1.57 [h,i] (**0.90**, 2.75)
Coggon et al. (2004)	1,012 female	TWA generally < 5 ppm; Peaks up to > 700 ppm	11	13.1	**0.84** [j] (0.42, 1.50)
Hogstedt et al. (1986)	153 female	TWA 20±10 ppm	0	---	**No breast cancer reported**
Meta-Analysis Studies					
Marsh et al. (2019) [k]					**0.97** (0.80, 1.18)
Vincent et al. (2019) [k]					**0.92** (0.84, 1.02)

[a] Based on external referent US population; see the text for information regarding why a healthy worker effect should not be expected for breast cancer incidence, an endpoint relied upon by USEPA (2016).

[b] Using the 233 cases with interviews as a surrogate, mean exposure level would be expected to be > 1 ppm since the mean is higher than the median in a lognormal distribution, median cumulative exposure for the 233 cases was 14.0 ppm-years, and mean years exposed was 13.0 (Table 2 of the study), so mean cumulative exposure >14 ppm-years/mean duration of 13 years = >1 ppm mean exposure.

[c] From Table 3 of the study based on workers whose exposure did not lag out using a 15-year lag period, consistent with USEPA (2016) and TCEQ; expected (E) value of 258.4 was calculated (i.e., E=O/0.89).

[d] For a 15-year lag, consistent with that used by USEPA (2016) and TCEQ.

[e] Breast cancer did not show any overall excess, although there was an excess in the highest cumulative exposure quartile (>12,322 ppm-days) using a 20-year lag and internal exposure-response analyses found positive trend for

breast cancer using the log of cumulative exposure with a 20-year lag but not cumulative exposure (Tables 1, 5, and 8 of study).

[f] From Table 3 of Mikoczy et al. (2011) and includes induction latency period of ≥15 years, consistent with that used by USEPA (2016) and TCEQ.

[g] This statistically significantly decreased breast cancer risk occurred in female workers exposed to a mean of ≈20 ppb EtO; this inordinately decreased SIR for the lowest cumulative exposure group produced statistically increased SIRs for higher cumulative exposure groups which did not experience increased breast cancer risk compared to the general population despite EtO mean exposures up to ≈1,110 ppb and more robust female worker data suggest that it represents an anomalous study artifact.

[h] For the most appropriate method identified by the study authors (Method 2) for the longest follow-up period (through 1987) with the most appropriate/matching SEER rates (through 1987) used to calculate the expected number (E).

[i] Includes two breast cancers diagnosed within 1 month of employment; reasonably excluding these two breast cancers diagnosed within 1 month of beginning work would not be expected to significantly reduce person-years but would result in a lower and still statistically insignificant estimated O/E (e.g., 10/7.64 = 1.31).

[j] For female workers with known continuous workplace exposure, the breast cancer mortality SMR was 0.70 (5 observed vs. 7.2 expected).

[k] This meta-analysis included all the individual studies above except for Hogstedt et al. (1986), which found no breast cancers and therefore did not report any effect estimate for breast cancer.

Given that EtO is purported by USEPA (2016) to be a potent breast cancer carcinogen, it is truly remarkable that a collectively large group of workers (e.g., the NIOSH cohort was 55% female, this study alone representing many thousands of workers) has been exposed to daily air concentrations up to tens-of-millions of times higher than typical environmental levels (as cited in USEPA 2016), yet a consistent and clear increase in breast cancer risk in the exposed is lacking.

A6.2 Healthy Worker Effect and Under-Ascertainment Considerations

The rationale behind epidemiological analyses with internal referents also requires evaluation in the present case, given that some internal analyses (based on a relatively small internal referent population) appear to show elevated breast cancer risk among EtO-exposed workers. Mikoczy et al. (2011) is a case in point. While study authors suggest that a healthy worker effect was indicated by significantly decreased overall mortality and cardiovascular disease mortality, *this cannot be assumed to necessarily extend to the incidence of a specific cancer*, particularly where the carcinogen operates via a mutagenic MOA (e.g., EtO). For example, the suggestion of the authors of Mikoczy et al. (2011) that a finding of significantly decreased overall mortality and cardiovascular disease mortality is indicative of a healthy worker effect for *breast cancer incidence* is inconsistent with the results of a relatively recent and large study (366,114 workers) conducted specifically to examine the potential for the healthy worker effect in cancer incidence studies (Kirkeleit et al. 2013). In Kirkeleit et al. (2013), all-cause mortality and both ischemic heart disease and circulatory system disease mortality were statistically significantly decreased in male workers (n=283,002) and female workers (n=83,112) compared to the

general population (Table 3 of the study), consistent with similar findings in Mikoczy et al. (2011). The SIRs for lymphoid and hematopoietic cancers in male workers and female workers were 0.97 (0.90, 1.03) and 1.09 (0.92, 1.27), respectively, consistent with the lack of a statistical difference as in Mikoczy et al. (i.e., SIR of 1.35 (0.54, 2.78) for lymphohaematopoietic cancer; Table 5 of the study). However, *most importantly and contrary to the implication by the authors of Mikoczy et al. (2011) that there is a health worker effect for breast cancer incidence* (requiring internal analysis for 1,197 female workers), *the Kirkeleit et al. (2013) study that specifically evaluated the potential for the healthy worker effect found that breast cancer incidence in over 83,000 female workers was as expected based on the general population* (i.e., SIR of 1.02 (0.95, 1.09)). *This strongly supports that the breast cancer SIR of 0.52 in Mikoczy et al. (2011) is an anomalous study artifact that should not be used for internal analyses*. Accordingly, this study shows no reliable evidence of increased breast cancer risk due to occupational EtO exposure. Again, the weight of evidence is that SIRs/SMRs across studies are not statistically significantly elevated for breast cancer (Table 42).

Similarly, for other studies, a presumption of the presence of a healthy worker effect for breast cancer incidence does not appear to be a robustly supported justification for internal analyses, which have the potential to use less reliable/stable referent rates based on much smaller worker populations than that used in Kirkeleit et al. (2013). Like Mikoczy et al. (2011), internal analyses conducted in Steenland et al. (2003) cannot be justified by the presumption of a healthy worker effect for breast cancer incidence. Instead, Steenland et al. indicate [*emphasis added*], "Because of the issue of *under-ascertainment*, we have emphasized internal exposure-response analyses in our study rather than the use of external referent population." Internal analyses of breast cancer incidence for cases with interviews were relied on by USEPA (2016). However, neither the healthy worker effect nor under-ascertainment justified internal analyses for breast cancer cases with interviews, as the study authors indicate [*emphasis added*], "Breast cancer *ascertainment* in the sub-cohort with interviews *was considered complete*."

Steenland et al. acknowledge that they found no excess of breast cancer incidence among the cohort as a whole compared to the US population; only finding an increase in the *highest exposure quintile* in certain *internal analyses*: that is, categorical with exposure lagged 15 years for cumulative exposure and duration exposure (see Tables 4 and 5 of Steenland et al. 2003). However, *without any justification for internal analyses in this case (as discussed above),* it is noted that using the external referent: (1) *The RR for even the highest exposed group* (>14,620 ppm-days) was *not statistically increased* (i.e., 1.27 (0.94, 1.69)) and *the RRs for all lower exposure groups were < 1*, consistent with no excess risk (see Table 3 of Steenland et al. 2003); and (2) *The overall RR for breast cancer incidence was 0.89* (0.78, 1.01) (see Table 42 above), *indicative of no excess risk overall* among 7,476 women workers with relatively high exposure to EtO. *Thus, considering that internal analyses appear unjustified in this case, no association of EtO with increased risk is demonstrated for the cohort overall or for any exposure category*.

In summary, consistent with the dubious biological plausibility of modeled risk results (Section 3.4.1.5.2.2) and USEPA's acknowledgement that human data alone are inadequate to classify EtO as a human breast carcinogen, the information discussed above confirms the dubious nature of epidemiological evidence of EtO-induced breast cancer. The recent meta-analysis by Marsh et al. (2019) of EtO studies that evaluated breast cancer reported a meta-RR of 0.97 (0.80, 1.18), leading study authors to conclude, "Evaluations of workers exposed during sterilization processes do not support the conclusion that EO exposure is associated with an increased risk of breast cancer". Similarly, the recent meta-analysis by Vincent et al. (2019) of EtO studies that evaluated breast cancer reported a meta-RR of 0.92 (0.84, 1.02). The study authors concluded, "Higher quality epidemiological studies demonstrated no increased risk of breast cancers."

A6.3 Relevance of Laboratory Animal Data

USEPA (2016) acknowledges that human data are insufficient to establish that EtO is a human breast cancer carcinogen, which again, would be quite unexpected if EtO were in fact as highly potent of a carcinogen as USEPA (2016) purports given the large group of workers (including women) exposed to very high concentrations of EtO on a daily basis. As a result, USEPA must rely on support from laboratory animal studies in classifying EtO as *carcinogenic to humans*. However, upon closer scientific scrutiny, the sites of EtO-induced cancers in animal models are of questionable direct human relevance for being predictive of, and therefore being used as confirming evidence for, the site(s) of human cancers.

While laboratory animal data are often used to support various aspects of regulatory assessments, interspecies differences in carcinogenic responses are common (e.g., tumors types, sensitivity), even between rodents (e.g., EtO-induced mammary tumors in mice but not rats). For example, IARC (2019) analyzed tumor site concordance using a dataset of the 111 distinct Group 1 (*carcinogenic to humans*) agents identified up to and including Volume 109. Sixty agents had both a human tumor site and an animal tumor site identified and were used to evaluate concordance across 39 tumor sites in animals and humans (see Figures 21.1 and 21.2 of IARC 2019). Reported results show that breast cancer is more frequently/commonly induced in laboratory animal species by these agents than in humans. More telling is that while there is 47% overlap between agents that cause lymphoid and haematopoietic cancers in humans and animals, there is only 20% overlap between agents that have been shown to cause breast cancer in humans and animals (Table 21.7 of IARC 2019). The IARC (2019) unanimous consensus statement is that "At present, the state of the science does not support tumour site concordance as a general principle."

Accordingly, animal data are not deterministic as to the sites of chemically-attributable carcinogenesis in humans; even more so when laboratory animal results are inconsistent (i.e., mammary tumors in mice but not rats) and the human database is relatively robust. For

example, *lung cancer* was statistically increased in both male and female mice at incidences of 53% and 45%, respectively (Table 3-3 in USEPA 2016), but is a not a candidate endpoint in humans as *data for this very strong carcinogenic response in mice is simply not predictive for humans* (i.e., no interspecies site concordance; SMR of 1.05 (0.95, 1.17) in Table 1 of Steenland et al. 2004). Similarly, EtO induced statistically significant increases in *brain tumors* in rats of both sexes (Table 3-5 in USEPA 2016), *but yet again, these results are not predictive for humans*. In fact, *brain cancer for the NIOSH cohort is statistically significantly decreased* (i.e., SMR of 0.59 (0.36, 0.91) in Table 1 of Steenland et al. 2004), just the *opposite* of what the rat data would suggest. Clearly, laboratory animal data for EtO-induced cancers cannot be relied upon to identify cancer sites or otherwise predict EtO carcinogenic response in humans. Thus, *the laboratory animal data are (to say the least) of dubious relevance for confirmation of, or adequately supporting, the inadequate epidemiological evidence of breast cancer as a known site of EtO-induced carcinogenesis in humans* (see above). In addition to evaluating epidemiological evidence, Vincent et al. (2019) also evaluated the animal study results, concluding that they provide no strong indication that EtO causes mammary tumors.

Based on the above considerations regarding the inadequate epidemiological evidence of EtO-induced breast cancer in humans and the dubious relevance and utility of laboratory animal data, USEPA's *carcinogenic to humans* classification is best supported by the lymphoid cancer data (e.g., see Figures 1 versus 4 of Vincent et al. 2019) and TCEQ's final URF is best based on lymphoid cancer as the critical cancer endpoint.

A6.4 Ecological Information of Interest

Although ecological information is associated with significant uncertainties, the public has expressed interest in such studies through public comment. As breast cancer is the subject of this appendix and one of the two cancer endpoints used by USEPA (2016), the TCEQ notes the mixed results:

- Breast cancer was not increased around Terumo BCT in Lakewood, CO, with an SIR of 0.98 (95% CI of 0.68, 1.35) (https://drive.google.com/file/d/1WEe0kCfkXW2RQC4jRFsIC803u_6P1Mub/view);
- Based on state *but not* county referents, breast cancer was statistically increased around Sterigenics in Willowbrook, IL (e.g., SIR of 1.1 with 95% CI of 1.02, 1.18) (http://www.dph.illinois.gov/sites/default/files/publications/sterigenicswillowbrookcancer-investigation-final.pdf); and
- Breast cancer was statistically significantly *decreased* around the Viant Medical Facility in Grand Rapids, MI, using both county rates (SIR of 0.81 with 95% CI of 0.71, 0.91) and state rates (SIR of 0.88 with 95% CI of 0.77, 0.99)

(https://www.michigan.gov/documents/mdhhs/Viant_Cancer_Incidence_Review_661354_7.pdf).

The TCEQ conducted a brief exploratory ecological analysis based on Texas and county EtO emissions data from the 2014 National Emissions Inventory (NEI; https://www.epa.gov/air-emissions-inventories/2014-national-emissions-inventory-nei-data). Based on the NEI, Texas emits approximately 36% of the EtO in the US. As a result, tons of EtO emitted per square mile in Texas (1.8E-04 tons/square mile) is over 5 times higher than the rest of the US as a whole (3.5E-05 tons/square mile). Despite this and the extraordinarily high carcinogenic potency purported by USEPA (2016) for lymphoid and breast cancers, *Texas incidence rates are statistically significantly lower than the US for non-Hodgkin's lymphoma, breast cancer, and all cancer sites combined, with leukemia being of borderline statistical significance (Table 43)*. Leukemia, non-Hodgkin's lymphoma, and breast cancer are endpoints included in USEPA's carcinogenic dose-response assessment for EtO (USEPA 2016), along with multiple myeloma (state-specific versus US data were not available).

Table 43: Some Example Age-Adjusted Cancer Incidence Rates per 100,000 (2012-1016)

Area	NEI Emissions (tons)	Emissions per Area (tons/ mile^2)	Non-Hodgkin's Lymphoma	Leukemia	Breast Cancer (female)	All Cancer Sites
US	133.72	3.52E-05	19.2 (19.1, 19.3)	14.1 (14.1, 14.2)	125.2 (124.9, 125.4)	448.0 (447.7, 448.4)
Texas	48.45	1.80E-04	17.4 (17.2, 17.6)	13.9 (13.7, 14.1)	111.9 (111.2, 112.7)	407.7 (406.6, 408.9)
Jefferson County	12.05	1.08E-02	17.5 (15.3, 19.9)	13.4 (11.5, 15.5)	102.4 (94.9, 110.3)	399.9 (389.3, 410.7)
Harris County	11.75	6.60E-03	16.9 (16.3, 17.5)	13.0 (12.5, 13.5)	111.9 (109.9, 114.0)	400.1 (397.2, 403.1)

As to breast cancer specifically, USEPA's assessment of EtO as a potent carcinogen suggests that elevations in EtO-induced breast cancers should be expected in counties with relatively high EtO emissions per square mile (as a surrogate for exposure) and a sufficiently large population. Figure 23 shows breast cancer incidence rates for the Texas counties with the highest EtO emissions per the NEI.

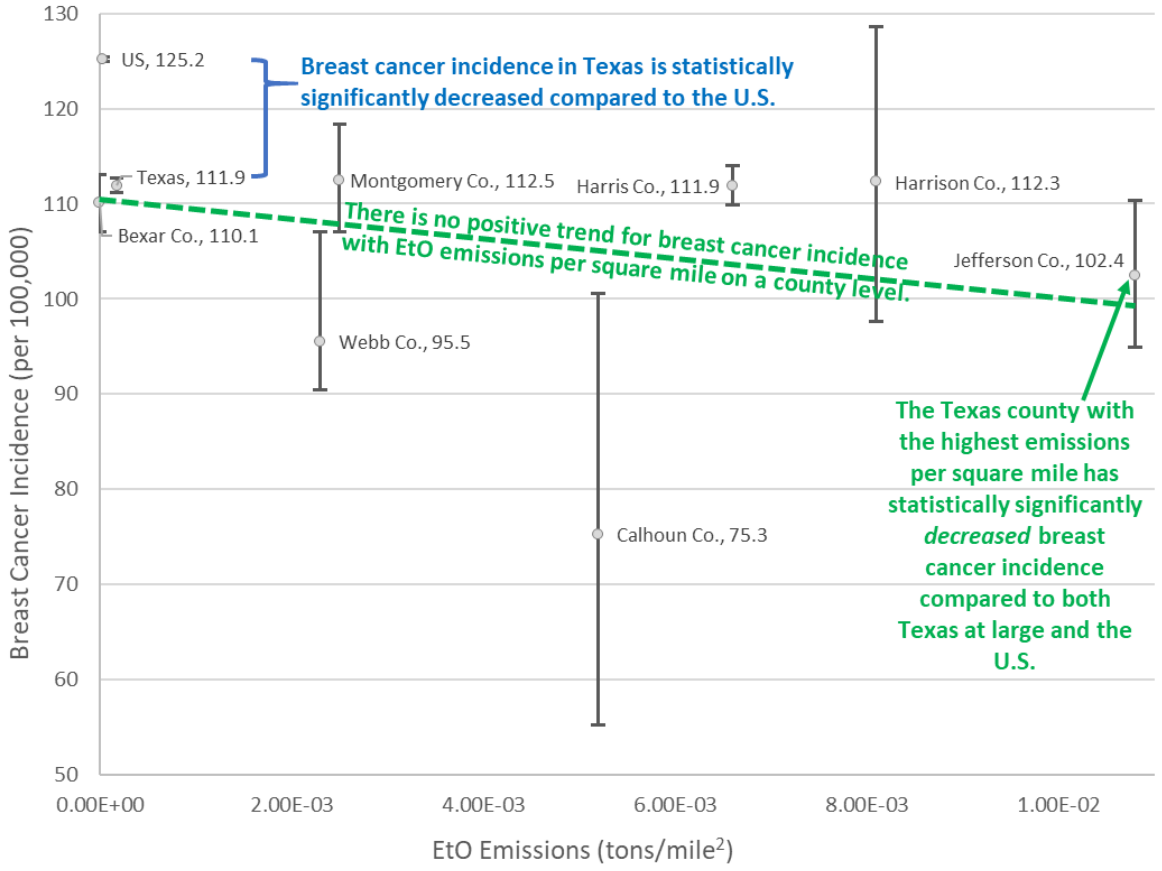


Figure 23: Texas and U.S. Breast Cancer Incidence versus NEI EtO Emissions per Square Mile on a County Level

The TCEQ notes although highly-industrialized Jefferson County (population ≈260,000) has more EtO emissions on a square mile basis than any other county in Texas (1.1E-02 tons/square mile) with over 300 times more than the US at large (3.5E-05 tons/square mile), the incidence of breast cancer (102.4 per 100,000 with 95% CI of 94.9, 110.3) is lower in Jefferson County, Texas than in the general US population (Figure 23, Table 43). *In fact, breast cancer incidence is statistically significantly lower in Jefferson County compared to both Texas and the US, despite EtO emissions that are 60 times higher than Texas at large and 307 times higher than the US*. Based on USEPA's 2016 assessment, the opposite of this reality would be expected. Similarly, as by far the most populated Texas county (≈4.6 million) with relatively high reported NEI EtO emissions per square mile (i.e., 6.6E-03 tons/square mile is ≈188 times higher than the US at 3.5E-05 tons/square mile), *the incidence of breast cancer (111.9.4 per 100,000 with 95% CI of 109.9, 114.0) is statistically significantly lower in highly-industrialized Harris County than in the general US population* (Figure 23, Table 43). Despite the associated uncertainties, such results may be viewed as surprising when considered in the context of USEPA (2016).

Appendix 7 PODs within the Observable Range of Key Cohort Data

For this DSD, the TCEQ evaluates the lower limit on the effective concentration (LEC; 95% LCL) at an extra risk of 1 in a 100,000 consistent with USEPA cancer guidelines (2005a) on the selection of a POD at the low-end of the observable range of exposures. Although for animal studies, a typical POD is an extra risk of 0.10 because it corresponds to doses near the low end of the doses, in epidemiological studies a lower level of risk often needs to be used.

The TCEQ uses the standard Cox proportional hazards model to calculate LEC for an extra risk of 1 in a 100,000 because the effective concentration (EC) corresponding to the same risk level are in the range of the observed data in the NIOSH study. That is, the EC for an extra risk of 1 in 100,000 of lymphoid cancer mortality in males is 9.67E-03 ppm for 70 years with an exposure lag of 15 years, which correspond to a cumulative occupational exposure of 591 ppm-days. There are 7 male workers in the NIOSH cohort with cumulative exposures less than 591 ppm-days. That is, 25.9% of the male workers in the NIOSH cohort that died with lymphoid cancer were exposed to cumulative exposures of less than the EC for 1 in a 100,000. In contrast, the EC for 1 in 100 results in environmental concentrations corresponding to cumulative occupational exposures of 354,400 ppm-days, which exceeds the largest cumulative exposure of lymphoid male decedents in the NIOSH study.

Table 44 shows the EC corresponding to different risk levels and the corresponding cumulative exposures with the number of lymphoid mortality cases of the male workers in the NIOSH study.

Table 44: Environmental and equivalent occupational cumulative EtO exposures for different potential PODs using TCEQ's selected model for lymphoid mortality in the NIOSH study (male workers)

Statistic	Extra Risk			
	1/100	1/1,000	1/10,000	1/100,000
Environmental EC (ppm) [1]	5.80×10^{-0}	8.99×10^{-1}	9.61×10^{-2}	9.67×10^{-3}
Equivalent Occupational EC (ppm-days) [2]	354,399	54,932	5,872	591
Lymphoid Deaths [3]	27	21	13	7
% Lymphoid Deaths [4]	100%	77.78%	48.15%	25.93%
% Male Workers [5]	99.84%	94.48%	66.45%	30.17%
LEC (ppm) [6]	2.44×10^{-0}	3.78×10^{-1}	4.04×10^{-2}	4.07×10^{-3}
URF (ppb^{-1}) [7]	4.09×10^{-6}	2.64×10^{-6}	2.47×10^{-6}	2.46×10^{-6}

[1] Environmental concentration in ppm for 70-year lifetime with lag of 15 years corresponding to a specified extra risk
[2] Equivalent Occupational Exposure 70 years (ppm-days) = EC (ppm) × (365/240 days) × (20/10 m^3) × (365.25 days/year) × (70 years – lag in years)
[3] Number of male workers in the NIOSH cohort that died of lymphoid cancer with cumulative exposure less than the EC (i.e., EC in ppm-days at 1/100, 1/1,000, 1,10,000, or 1/100,000)
[4] Percentage of lymphoid cancer decedent male workers in the NIOSH cohort with cumulative exposures less than the EC (ppm-days)
[5] Percentage of male workers in the NIOSH cohort with cumulative exposures less than the EC (ppm-days)
[6] 95% lower bound on the EC (ppm)
[7] Unit risk estimate based on the LEC (ppm)

The results in Table 44 show that the EC for an extra risk of 1 in a 100 is outside the range of cumulative exposures for the male lymphoid mortalities observed in the NIOSH study and in the upper 1% of cumulative exposures for all male workers. That is, all males that died with lymphoid cancers and more than 99% of all male workers had cumulative exposures less than the EC (1/100). Thus, the NIOSH study does not support an extra risk of 1 in a 100 as a POD.

The EC for an extra risk of 1 in a 1,000 is a concentration that is at the high-end of cumulative exposures of male lymphoid mortalities observed in the NIOSH study. That is, 77.78% of all males that died with lymphoid cancers and 94.48% of all male workers had cumulative exposures less than the EC (1/1,000). Thus, a POD of 1 in 1,000 is at the higher-end of the cumulative exposures of male workers of the NIOSH study.

The EC for an extra risk of 1 in 10,000 is a concentration that includes 48.15% of the decedent men with lymphoid cancer and 66.45% of all men in the NIOSH cohort with smaller cumulative exposures. The EC for an extra risk of 1 in 100,000 includes 25.93% of male lymphoid decedents and 30.17% of all males in the NIOSH study with smaller cumulative exposures. Thus, use of an extra risk of 1 in 100,000 is supported by the NIOSH observed data, being near the lower end of the observed range of cumulative exposures to EtO, and is consistent with USEPA and TCEQ guidelines (USEPA 2005a, TCEQ 2015) on the selection of a POD at the low-end of the observable range of exposures.

Based on Table 44 results, using either 1 in 10,000 or 1 in 100,000 extra risk PODs (as PODs in the range of the observed data and close to the low-end of the observable range) round to the same ADAF-unadjusted URF selected by the TCEQ (2.5E-06 per ppb; Table 13). Looking at it from a different perspective, using the 1 in 10,000 excess risk LEC of 4.04E-02 ppm as the POD and linear extrapolation, the 1 in 100,000 air concentration (ADAF unadjusted) is still 4 ppb (i.e., 1E-05/2.47E-06 per ppb = 4.05 ppb).