Aaron Szabo
Assistant Administrator
Office of Air and Radiation (OAR)
U.S. Environmental Protection Agency
1200 Pennsylvania Avenue NW Washington, DC 20460

May 15, 2026

Re: The Ethylene Oxide Sterilization Association's Comments on the *National Emission Standards for Hazardous Air Pollutants: Ethylene Oxide Emissions Standards for Sterilization Facilities Residual Risk and Technology Review Reconsideration*, 91 Fed. Reg. 12,700 (March 17, 2026); Docket EPA-HQ-OAR-2019-0178

Dear Assistant Administrator Szabo:

The Ethylene Oxide Sterilization Association ("EOSA") submits these comments in response to the U.S. Environmental Protection Agency's ("EPA's") proposed rule, *National Emission Standards for Hazardous Air Pollutants: Ethylene Oxide Emissions Standards for Sterilization Facilities Residual Risk and Technology Review Reconsideration,* 91 Fed. Reg. 12,700 (March 17, 2026) ("Proposed Rule"). EOSA greatly appreciates and supports the EPA's reconsideration of the 2024 Rule,[1] which was legally flawed and set technically unachievable ethylene oxide ("EtO") emission standards for sterilizers, threatening the U.S. medical device supply chain.

EOSA is comprised of the leading experts of the medical device manufacturing and sterilization industry. EOSA members include ethylene oxide sterilizers, medical device manufacturers, emissions control equipment manufacturers, and ethylene oxide detection systems manufacturers. EOSA members also include contract sterilization providers, sterilization consultants, environmental consultants, and registrants. Many EOSA members are directly regulated by the 2024 Rule, and they faced difficult choices as they considered whether it was possible to comply with the extremely stringent emission limitations and monitoring requirements set in that Rule.

EOSA strongly supports the EPA's Proposed Rule and urges the EPA to finalize it substantially as proposed. EPA is right to recognize that parts of the 2024 Rule— including all emission standards promulgated under Clean Air Act ("CAA") section 112(f)(2)—were unlawful, while others (e.g., the requirement that sterilizers demonstrate compliance only through continuous emissions monitoring systems ("CEMS")) were technically infeasible. Some aspects of the 2024 Rule, such as the EPA's reliance on the flawed 2016 Integrated Risk Information System ("IRIS") risk value for EtO to promulgate emission standards, were both legally and technically unsound. EOSA urges the EPA to finalize the Proposed Rule to address those errors in the 2024 Rule and reinstate a sensible regulatory framework that will enable EOSA's members to continue operating responsibly to supply sterilized medical devices to U.S. healthcare providers while limiting emissions of EtO based on technology that can be efficiently and cost-effectively deployed.

---

[1] *National Emission Standards for Hazardous Air Pollutants (NESHAP): Ethylene Oxide Emissions Standards for Sterilization Facilities Residual Risk and Technology Review*, 89 FR 24,090, April 5, 2024 ("2024 Rule").

EOSA's comments in response to the EPA's questions, as well as some additional issues that EOSA raises for the EPA's consideration, are divided into three volumes:

Volume 1 addresses key legal issues in response to the EPA's questions #1-2, including that the text and structure of CAA section 112 do not allow the EPA to conduct multiple rounds of risk review under section 112(f)(2). Volume 1 also addresses the legal errors in the EPA's heavy reliance on the 2016 IRIS in the 2024 Final Rule and other cross-cutting legal issues implicated by questions #3-6.

Volume 2 addresses the technical side of the EPA's questions #3-6, which invite reevaluation of the 2016 IRIS value and call for additional data and studies on risk from EtO exposure. EOSA submits a "reader's guide" to these technical comments, followed by a comprehensive analysis of the relevant studies and data, demonstrating that the 2016 IRIS value does not reflect best available science.

Volume 3 responds to the EPA's questions #7-21, addressing technical issues raised by the Proposed Rule. Key among these are the EPA's proposed rescission of the Permanent Total Enclosure ("PTE") mandate, and the EPA's proposal to allow facilities to demonstrate compliance with standards through either parametric monitoring or continuous emissions monitoring systems ("CEMS"). EOSA also includes, in Volume 3, additional technical comments for the EPA's consideration, such as that the EPA should promulgate alternative outlet concentration limits to enable compliance by highly efficient control systems operating with low inlet concentrations.

As outlined above, EOSA's comments cover a broad range of legal and technical issues relevant to EPA's reconsideration of the 2024 Final Rule, and EOSA hopes that the EPA will find all of the information it has provided helpful in expeditiously finalizing its Proposed Reconsideration Rule. But there are certain issues that EOSA's members view as critical to their ability to continue to provide a stable and robust supply of sterilized medical devices. It is most important that the EPA:

- Rescind the stringent risk-based standards unlawfully set in the 2024 Rule;
- Cease relying on the 2016 IRIS to promulgate emission standards under section 112;
- Rescind the PTE mandate and provide for enhanced compliance flexibility to demonstrate capture efficiency using EPA Method 204;
- Allow sterilizers to demonstrate compliance through *either* CEMS *or* parametric monitoring;
- Promulgate alternative outlet concentration limits alongside the DRE-form standards; and
- Provide three years for regulated facilities to meet the revised standards and requirements.

If the EPA takes the above actions and approaches in finalizing the Proposed Rule, EOSA's members anticipate that they will be able to continue to meet U.S. demand for sterilized medical devices and other essential sterilized products and equipment, while controlling emissions in a technically feasible and cost-effective way to protect public health.

EOSA again thanks the EPA for reconsidering the legally and technically flawed 2024 Final Rule and urges the EPA to finalize the reconsideration rule with the modifications discussed above expeditiously as it can while still considering all relevant information.

Sincerely,

Meibao Zhuang

Meibao Zhuang
Senior Manager
The Ethylene Oxide Sterilization Association, Inc.

BEFORE THE
UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
OFFICE OF AIR AND RADIATION

COMMENTS OF THE ETHYLENE OXIDE STERILIZATION ASSOCIATION ON OAR'S PROPOSED NATIONAL EMISSION STANDARDS FOR HAZARDOUS AIR POLLUTANTS: ETHYLENE OXIDE EMISSIONS STANDARDS FOR STERILIZATION FACILITIES RESIDUAL RISK AND TECHNOLOGY REVIEW RECONSIDERATION

Air Docket (EPA-HQ-OAR-2019-0178) Opened for Review and Comment (91 Fed. Reg. 12700 (March 17, 2026)	)	EPA-HQ-OAR-2019-0178

Ethylene Oxide Sterilization Association
2200 Pennsylvania Avenue, N.W.
Suite 100W
Washington, DC 20037

Introduction

The Ethylene Oxide Sterilization Association (EOSA) submits these comments to the United States Environmental Protection Agency (EPA) in response to the March 17, 2026, notice of availability of *the Proposed National Emission Standards for Hazardous Air Pollutants: Ethylene Oxide Emissions Standards for Sterilization Facilities Residual Risk and Technology Review Reconsideration* (Proposed Reconsideration of Sterilizer NESHAP, or Proposed Reconsideration).[1]

EOSA is comprised of the leading experts of the medical device manufacturing and sterilization industry. EOSA members include ethylene oxide sterilizers, medical device manufacturers, manufacturers of emissions control equipment and ethylene oxide detection systems. EOSA members also include many contract sterilization providers, sterilization consultants, environmental consultants, and registrants. The diversity of EOSA member companies provides EOSA a broad view of the overall implications of these proposed regulatory actions. EOSA's members are uniquely capable of assessing the issues and complications posed by the proposed reconsideration of the sterilizer NESHAP on the sterilization process, the sterilization facilities, the economics of sterilization, the safety of sterilization facilities, and the security of the supply chain of critical medical devices.

EOSA appreciates the opportunity to provide comments on EPA's proposal to revise the 2024 Sterilizer Rule,[2] which is of utmost importance to EOSA's members. The 2024 Rule imposed extremely stringent standards and requirements that many sterilizers likely would not have been able to meet, causing some to cease operations permanently. Others would have had to shut down temporarily to reconfigure their facilities, interrupting the supply of sterilized medical devices. EOSA therefore strongly supports the Proposed Reconsideration Rule, with some suggested changes to certain standards and requirements and additions to the EPA's legal reasoning and technical analysis. The critical areas of concern for EOSA's members are:

- **Rescission of Risk-Based Standards:** We strongly support the EPA's determination that the 2024 Final Rule relied on an unauthorized second residual risk review under CAA section 112(f)(2). The EPA is correct to recognize, in the Proposed Reconsideration, that the statutory text and scheme do not allow for

---

[1] 91 Fed. Reg. 12700 (March 17, 2026) (*Federal Register* notice announcing the availability of the U.S. Environmental Protection Agency's (EPA) Proposed National Emission Standards for Hazardous Air Pollutants: Ethylene Oxide Emission Standards for Sterilization Facilities Residual Risk and Technology Review Reconsideration (Proposed Reconsideration of Sterilizer NESHAP of Proposed Reconsideration), available at *https://www.federalregister.gov/documents/2026/03/17/2026-05167/national-emission-standards-for-hazardous-air-pollutants-ethylene-oxide-emissions-standards-for*.

[2] *National Emission Standards for Hazardous Air Pollutants (NESHAP): Ethylene Oxide Emissions Standards for Sterilization Facilities Residual Risk and Technology Review*, 89 FR 24090, April 5, 2024 ("2024 Rule").

multiple rounds of risk review; standards may be revised if appropriate pursuant to the periodic technology-based review under section 112(d) after balancing costs and benefits as mandated by Congress. We accordingly encourage the EPA to finalize the rescission of the risk-based standards set in the 2024 Final Rule.

- **Reliance on the 2016 EtO Integrated Risk Information System (IRIS) Value:** We urge the EPA not to rely on the 2016 IRIS value when setting standards under CAA section 112, which does not reflect best available science. Reliance on the 2016 IRIS leads to risk estimates that are often indistinguishable from natural background levels, imposing regulatory burdens that do not align with real-world health outcomes. As intended by Congress, the EPA should consider a broad range of relevant studies and data addressing EtO risk and carcinogenicity, which EOSA identifies and discusses in Volume 2.

- **Technical Feasibility and PTE:** We support the EPA's proposal to rescind the mandate for Permanent Total Enclosures (PTE), which remains technically prohibitive for many existing facility configurations and lacks a clear pathway for practical implementation. Where capture efficiency must still be demonstrated, we strongly urge the Agency to provide enhanced compliance flexibility under the EPA Method 204 to account for the unique operational realities, airflow dynamics, and layout constraints of existing sterilization facilities.

- **DRE and Alternative Outlet Concentration Limits:** While we recognize the Agency's goal of maximizing emissions reductions, strict adherence to a high Destruction and Removal Efficiency (DRE) is often technically unachievable for control systems operating with low inlet concentrations due to the physical detection limits of monitoring instruments. We strongly advocate for the inclusion of an alternative outlet concentration limit to provide a scientifically sound and practically achievable pathway for highly efficient control systems operating under these very low EtO concentrations.

- **Monitoring Flexibility and CEMS Alternatives:** We strongly support the EPA's proposal to allow facilities the option to utilize parametric monitoring in lieu of Continuous Emissions Monitoring Systems (CEMS). The rigid mandate for CEMS under Performance Specification 19 introduces significant operational burdens, matrix interference challenges, and disproportionate costs, particularly at ultra-low detection levels. To maximize this flexibility, we urge the Agency to ensure the final rule does not lock facilities into rigidly prescribed parameters or limited technological options, but instead to permit a site-specific, adaptable approach to parametric monitoring.

- **Medical Supply Chain Integrity:** Above all, these comments emphasize the need for a stable, balanced regulatory framework that protects the air we breathe without jeopardizing the domestic supply chain for the 20 billion medical devices sterilized annually with EtO in the United States.

To assist the EPA in the finalization of the Commercial Sterilizer NESHAP, EOSA's comprehensive comments have been organized into three volumes:

- **Volume 1 (Legal Issues):** Responds to the threshold legal questions posed by EPA Questions 1 and 2 and addresses certain cross-cutting legal issues relevant to other questions, such as whether EPA should rely on the EtO 2016 IRIS or must consider other data and studies in order to comply with the CAA and ensure all emission standards are reasonable and supported by the record.

- **Volume 2 (Scientific Data on EtO Risk):** Addresses EPA questions 3-6 by providing information about and analysis of additional data and studies that EPA should consider when assessing risk related to sterilizer EtO emissions. EOSA includes, as Volume 2(A) at the start of this set of comments, a "Reader's Guide" to the comprehensive scientific analysis set forth in Volume 2(B).

- **Volume 3 (Technical Comments):** Responds to EPA questions 7-21 and provides additional technical comments to ensure operational feasibly and address other considerations required for practical implementation of the standards proposed by the EPA in the Proposed Reconsideration Rule.

Volume 1

Comments on Legal Issues

EOSA appreciates the opportunity to provide comments on the EPA's proposal to revise the 2024 Sterilizer Rule, including to rescind standards based on a legally impermissible second round of risk review. In this first volume of its comments, EOSA responds to the threshold legal questions posed by the EPA and addresses certain cross-cutting legal issues relevant to other questions, such as whether EPA should rely on the EtO IRIS or consider other data and studies.

Responses to many of the EPA's technically focused questions are addressed in Volume 2 of EOSA's comments. Volume 3 of EOSA's comments addresses questions 3-6 by providing information about and analysis of additional data and studies that the EPA should consider when assessing risk related to sterilizer EtO emissions.

Question #1: Has the 2024 Final Rule and underlying interpretations generated reliance interests and, if so, how should the EPA consider them in any final action?

In response to the EPA's question 1, we provide the specific comment below.

Stakeholders have no reliance interests in the Final Rule because it never took effect. Depending on the operative provision of the Clean Air Act ("CAA" or "the Act"), the Final Rule's standards were to take effect either two or three years after promulgation. For standards set under CAA section 112(f)(2), the effective date was April 6, 2026. *See* 89 Fed. Reg. 24,090, 24,101 (Apr. 5, 2024). For standards set under CAA section 112(d), the effective date was April 5, 2027. *See id.* at 24,102. As the EPA is aware, the agency committed to reconsidering the Final Rule in early 2025, *see* Proposed Reconsideration Rule, 91 Fed. Reg. 12,700, 12,705 (Mar. 17, 2026), well before any of the new standards became effective. And well before that, the standards were challenged by both environmental and industry groups in petitions for review filed in the D.C. Circuit (Nos. 24-1178 & 24-1180) shortly after promulgation of the Final Rule. In that litigation, industry petitioners argued that the EPA lacks authority under the Act to conduct a second round of risk review and set new standards under section 112(f) in addition to challenging many of the Rule's standards and requirements on technical grounds. Given that pending litigation coupled with EPA's commitment to reconsider the Final Rule, stakeholders cannot reasonably have relied on that Rule remaining in place as promulgated or taking effect.

Furthermore, the President issued a Proclamation in July 2025 exempting many sterilizer facilities from the obligation to comply with any of the Final Rule's standards or deadlines for *an additional two years* "from the date originally required for such [compliance] deadline." Proclamation No. 10959, 90 Fed. Reg. 34,747 (July 17, 2025). That extended the effective date for section 112(f) standards to April 2028, and for section 112(d) standards to April 2029.

Thus, neither the regulated industry, nor the public, nor any other stakeholder could have reasonably relied on the Final Rule's standards and requirements to act or make decisions. Indeed, sterilizers regulated under the Final Rule relied on the fact that the Rule's standards would *not* govern their facilities for several years—and most likely would be revised before they came into effect. Conversely, it is difficult to see what sort of reliance interest an unregulated party could have in the Rule remaining in place exactly as finalized in 2024. While some might

argue that there is a public interest in reducing emissions, that is equally true in all CAA rulemaking proceedings and cannot suffice to sustain a rule where there are legal and technical reasons it should be revised. Further, there is also a public interest in the EPA's compliance with the requirements and limitations imposed by Congress, which the EPA is advancing by proposing changes to bring the Rule into compliance with CAA section 112.

Finally, the EPA correctly suggests that, even if some stakeholders were to claim reliance interests, those interests cannot supersede the best reading of CAA section 112. *See* 91 Fed. Reg. at 12,705. Put differently, if the Final Rule is unlawful in the respects that the EPA has now recognized, then it cannot be maintained regardless of any reliance on it. That is particularly true for a Rule published after the Supreme Court's decision in *Loper Bright Enterprises v. Raimondo*, 603 U.S. 369 (2024), and that was never upheld by a court as consistent with—much less the best interpretation of—the governing statute. Although there is pre-*Loper* caselaw holding that the EPA must *consider* reliance interests (if there are any, which is not the case here) as part of the notice-and-comment process, even under that precedent it is "free to 'conclude that reliance interests in [standards] that it views as unlawful are entitled to no or diminished weight.'" *MediNatura, Inc. v. Food & Drug Admin*., 496 F. Supp. 3d 416, 456 (D.D.C. 2020), *aff'd*, 998 F.3d 931 (D.C. Cir. 2021) (quoting *Dep't of Homeland Sec. v. Regents of the Univ. of California*, 591 U.S. 1, 32 (2020)). This is because, "[w]hile *justifiable* reliance interests are an appropriate consideration when weighing the proper scope of agency action, parties 'cannot claim justifiable reliance or protectable expectations based on ... [agency] action which was illegal.'" *Williams Nat. Gas Co. v. FERC*, 943 F.2d 1320, 1343 (D.C. Cir. 1991) (Edwards, J., dissenting) (quoting *Tennessee Valley Mun. Gas Ass'n v. FPC*, 470 F.2d 446, 453 (D.C. Cir. 1972)); *see also A.C.R. v. Noem*, 809 F. Supp. 3d 103, 122 (E.D.N.Y. 2025) (quoting same).

Accordingly, no justifiable reliance interests were generated by the Final Rule. To the extent any stakeholders claim reliance interests based on an unlawful reading of CAA section 112, the EPA should conclude that those interests are entitled to little or no weight.

Question 2: Should the EPA rescind the standards set under CAA section 112(f)(2) in the 2024 Final Rule, as proposed, based on the proposed finding that the discretionary second risk review conducted as part of the 2024 Final Rule was not authorized under the CAA?

In response to the EPA's question 2, we provide the specific comment below.

Yes. EOSA agrees that the EPA lacked authority to conduct a second round of risk review under CAA section 112(f)(2) and to set or revise standards under that provision in the Final Rule. There are many reasons that EPA's conclusion reflects the "single, best meaning" of the statute under *Loper Bright Enters.* v. *Raimondo,* 603 U.S. 369, 400, 411 (2024). Although the EPA identifies some of those reasons in the Proposed Reconsideration Rule, there are others that the EPA does not include or only partially addresses. EOSA suggests that the EPA include these additional reasons in its Final Reconsideration Rule to supplement and strengthen its rationale.

1. Statutory Text and Structure

The EPA correctly explains that the text and structure of CAA section 112 compel the conclusion that the EPA may only undertake one round of residual risk review. 91 Fed. Reg. at

12,712, 12,713. Agency action requires explicit statutory authorization. *See Hikvision USA, Inc. v. FCC*, 97 F.4th 938, 944 (D.C. Cir. 2024). No such explicit authorization to conduct seriatim risk reviews and revise risk-based emission standards for hazardous pollutants was granted by Congress in section 112(f) or elsewhere in the CAA.

Under section 112(f), residual risk standards are a *one-time event* contingent on a specific sequence of prerequisites, including: a report submitted to Congress no later than 6 years after November 1990, *see* 42 U.S.C. § 7412(f)(1); a failure by Congress to act on that report, *see id.* § 7412(f)(2)(A); the promulgation of standards under section 112(d), *see id.*; and finally, 8 years following such promulgation, a residual risk review, *see id.* By contrast, Congress used explicit language to establish a *recurring* technology review "no less often than every 8 years" under section 112(d)(6). *Id.* § 7412(d)(6).

Section 112(d), the technology review provision, is not the only part of section 112 that explicitly either requires or allows review and revision of the EPA's standards, requirements, or other actions addressing HAP emissions. Subsection (c)(1) provides that the EPA "shall from time to time . . . no less often than every 8 years, revise, if appropriate" the list of HAP source categories and subcategories. 42 U.S.C. § 7412(c)(1). Similarly, subsection (r)(3) authorizes EPA to "revise[] from time to time" and "review[] at least every 5 years" the list of substances posing risk of death, injury, or adverse effects if accidentally released. *Id.* § 7412(r)(3). And subsection (i)(8)(c) grants the EPA the authority to "review the emission limitations promulgated under subparagraph (B) [addressing coke oven emissions] and revise, as necessary, such emission limitations" by a certain time. *Id.* § 7412(i)(8)(C). Indeed, as the court noted in *Citizens for Pennsylvania's Future v. Wheeler*, 469 F. Supp. 3d 920, 929 (N.D. Cal. 2020), "while section 7412(f) says nothing explicit about revising risk-based standards, neighboring subsections direct EPA in no uncertain terms to 'revise' four other types of regulations: the list of hazardous air pollutants, the list of source categories, technology-based emission standards, and certain emissions limitations. 42 U.S.C. § 7412(b)(2), (c)(1), (d)(6), (i)(8)(C)."

Congress's choice not to include such explicit language authorizing subsequent rounds of review and revision from the residual risk review provisions must be read as an intentional limitation of that review process to a single, contingent instance in time. *See Hamdan v. Rumsfeld*, 548 U.S. 557, 578 (2006) ("[A] negative inference may be drawn from the exclusion of language from one statutory provision that is included in other provisions of the same statute.") *Republic of Sudan v. Harrison*, 587 U.S. 1, 12 (2019) ("Congress generally acts intentionally when it uses particular language in one section of a statute but omits it in another."). Indeed, in applying this canon in precisely this context, courts have observed that "[w]hile section [1]12(f) says nothing explicit about revising risk-based standards, neighboring subsections direct the EPA in no uncertain terms to 'revise'" other types of regulations. *Citizens for Pennsylvania's Future v. Wheeler*, 469 F. Supp. 3d 920, 929 (N.D. Cal. 2020). "Thus, where Congress wanted to establish a recurring obligation in section [1]12, the statute appears to accomplish that goal by express language, not by latent implication." *Id.* Interpreting the statute to imply authority for the EPA to conduct discretionary risk reviews on an ad hoc basis, as the EPA did in the 2024 Final Rule, undermines Congress's carefully constructed statutory framework and eliminate the finality of the distinct residual risk review process.

Other CAA programs confirm that Congress's choice not to explicitly require— or even allow—repeated risk reviews as part of the hazardous pollutant program under section 112 must be viewed a deliberate one:

A. The NAAQS program: Like the technology review provision of section 112(d), section 108 and 109 of the Act call for EPA to periodically review and revise standards under the National Ambient Air Quality Standards ("NAAQS") program. Section 108 directs the Administrator to "from time to time review, and, as appropriate, modify, and reissue any criteria or information on control techniques issued pursuant to this section," as well as to "revise and reissue criteria relating to concentrations of NO_2 . . . ." 42 U.S.C. § 7408(c). Section 109, in turn, directs the Administrator to, "at five-year intervals[,] . . . complete a thorough review of the criteria published under section 7408 of this title and the [NAAQS] promulgated under this section" and "make such revisions in such criteria and standards and promulgate such new standards as may be appropriate." *Id.* § 7409(d)(1). The use of clear language to require or allow periodic review and revision of standards set under other health-based CAA programs demonstrates that Congress knows how to grant the EPA such authority—and it would have included such language in section 112(f) had it intended to either require or allow periodic review and revision of risk-based HAP standards. *See Hamdan*, 548 U.S. at 578; *Republic of Sudan*, 587 U.S. at 12.

B. The NSPS program: In section 111 of the Act, Congress crafted a provision that instructs the EPA to periodically review and revise performance standards for new stationary sources, but also allows the EPA to find that such review/revision is not necessary. In subsection (b)(1), the statute directs the Administrator to, "at least every 8 years, review and, if appropriate, revise" standards for stationary sources. 42 U.S.C. § 7411(b)(1)(B). However, "the Administrator need not review any such standard if the Administrator determines that such review is not appropriate in light of readily available information on the efficacy of such standard." *Id*.; *see also id.* § 7411(g) (providing procedure for Governor of a State to apply for revision of standards with discretion for Administrator to grant or deny such application). Congress would have crafted similar language to include in Section 112(f) if it wanted to allow periodic risk review but was concerned that such review would only occasionally or infrequently be necessary.

All these textual and contextual aspects of the Act and Section 112 individually and together support the EPA's conclusion that Congress did not grant it authority to conduct additional rounds of risk review. Once the EPA concluded its one round of risk review and resultant standard-setting prescribed by section 112(f)(2)—as EPA did in 2006 for ethylene oxide emissions from sterilizers—the EPA's mandate under that section was concluded and its authority exhausted.

2. Legislative History

The legislative history of section 112 of the CAA also supports the EPA's proposed revised interpretation. EOSA notes that legislative history is a secondary statutory interpretation tool that cannot take precedence over analysis of a statute's text and structure. *See, e.g.*, *Cole v. Burns Int'l Sec. Servs*., 105 F.3d 1465, 1472 (D.C. Cir. 1997) ("where the statutory text does not admit of serious ambiguity, . . . legislative history is, at best, secondary, and, at worst, irrelevant"). Nevertheless, EOSA suggests that EPA recognize, in the Final Rule, that the history

of CAA section 112 confirms the EPA's proposed conclusion that the best reading of that provision is that EPA lacks authority to undertake multiple rounds of risk review.

As originally enacted in 1970, CAA section 112 required the EPA to set standards based on public health risk regardless of cost and technical feasibility, and the EPA came to recognize this mandate was impracticable. *See, e.g.*, Proposed Standard for Vinyl Chloride, 40 Fed. Reg. 59,532, 59,534 (Dec. 24, 1975) ("Complete prohibition of all emissions could require closure of an entire industry."). Thus, the first twenty years of section 112 amounted to "a lot of action, but nothing final." H.R. Rep. No. 101-490, at 151; *see also id*. ("No decision—is the history of this program."). The 1990 amendments transformed that stymied regulatory process by directing the EPA to promulgate standards while expressly considering cost and feasibility, and then review those standards every eight years to keep them in line with technological developments. *See* 42 U.S.C. § 7412(d)(6); *La. Env't Action Network*, 955 F.3d at 1093 ("That [technology] review ensures that, over time, the EPA maintains source standards compliant with the law and on pace with emerging developments[.]"). The residual risk review, on the other hand, was simply a fallback program incorporating preexisting language from before the amendments as a secondary safety net to the initial set of standards promulgated under the technology review process. *See* 136 Cong. Rec. S16,932, S16,979 (Oct. 27, 1990) ("We simply return to current law in the [residual risk] phase and ask EPA to set standards which provide an ample margin of safety to protect public health. That is the current law standard. It is also the reason that very little has been done to regulate toxic air pollutants[.]").

This history confirms that the residual risk review was intended as a one-time, gap-filler mechanism, not as a basis for recurring or ad hoc standard-setting that would circumvent the technology review process. Congress would not have gone through the trouble of creating a distinct, recurring technology review under section 112(d) if it intended the EPA to conduct open-ended risk reviews serving the same revisionary function. The EPA can therefore note, in the Final Reconsideration Rule, that the legislative history of section 112 reinforces what the statute's text and structure clearly show: Congress intended the EPA to conduct only one round of risk review, and make any subsequent revisions to standards pursuant to section 112(d).

3. Regulatory History and Case Law

The EPA's own regulatory history also supports its proposed revised interpretation. "[T]he longstanding practice of the government—like any other interpretive aid—can inform [a court's] determination of what the law is." *Loper Bright*, 603 U.S. at 386 (internal quotation marks omitted). From the enactment of section 112(f) in 1990 until 2024, the EPA did not rely on that provision to amend residual risk standards.

In 2012, for example, the EPA promulgated standards for the pulp-and-paper source category and explained that section 112(f)(2), "known as the residual risk review—is a one-time review that must occur within 8 years of issuance of the MACT standard." NESHAP From the Pulp and Paper Industry, 77 Fed. Reg. 55,698, 55,699 (Sep. 11, 2012). It reiterated this position in 2016 when it proposed amendments to standards for specific sources. *See* NESHAP for Chemical Recovery Combustion Sources at Kraft, Soda, Sulfite, and Stand-Alone Semichemical Pulp Mills (Proposed Rule), 81 Fed. Reg. 97,046, 97,048 (Dec. 30, 2016).

The EPA made the same statements in 2022 and 2023 in rulemakings for electric utility steam generating units: "CAA section 112(f)(2) requires the EPA to conduct *a one-time review* of the risks remaining after imposition of MACT standards under CAA section 112(d)(2) within 8 years of the effective date of those standards (risk review). CAA section 112(d)(6) requires the EPA to conduct a review of all CAA section 112(d) standards at least every 8 years . . . ." NESHAP: Coal- and Oil-Fired Electric Utility Steam Generating Units (Notice of Proposed Rulemaking), 87 Fed. Reg. 7,624, 7,632 n.18 (Feb. 9, 2022); Final Action, 88 Fed. Reg. 13,956, 13,962 n.22 (Mar. 6, 2023) (same).

In 2024, when finalizing amendments to these standards pursuant to a 2020 residual risk and technology review, the EPA again underscored the singularity of the risk review: "Congress required the EPA to conduct the CAA section 112(d)(6) review of what additional reductions may be obtained based on new technology *even after the EPA has conducted the one-time CAA section 112(f)(2) risk review* . . . . The two requirements are distinct . . . ." NESHAP: Coal- and Oil-Fired Electric Utility Steam Generating Units, Review of the Residual Risk and Technology Review, 89 Fed. Reg. 38508, 38,514 (May 7, 2024). Indeed, even in the 2024 Final Rule that the EPA is now proposing to reconsider, the EPA acknowledged: "CAA section [1]12(f)(2) requires only a *one-time risk review,* which is to be conducted within eight years of the date the initial standards are promulgated[.]" 89 Fed. Reg. at 24,094.

The EPA's long history of describing risk review under section 112(f) as a "one-time" event is consistent with how the D.C. Circuit has characterized that provision. In 2015, while discussing section 112's "two distinct" review processes, the D.C. Circuit described the section 112(f)(2) risk review as a "one-time" event and the section 112(d) technology review as "recurring." *Nat'l Ass'n for Surface Finishing v. EPA*, 795 F.3d 1, 5 (D.C. Cir. 2015); *see also La. Env't Action Network v. EPA*, 955 F.3d 1088, 1093 (D.C. Cir. 2020) ("section 112(d)(6)[] requires EPA, on an ongoing periodic basis, to revisit and update emission standards," and "EPA under section 112(f)(2) must conduct a one-time review").

The EPA's proposed interpretation of section 112 is therefore consistent with regulatory history and case law, whereas the 2024 Final Rule was at odds with both, warranting correction.

4. Cost/Benefit Analysis

In the Proposed Reconsideration Rule, the EPA does not address an important reason why it would be inconsistent with the statutory scheme and legislative history to engage in more than one round of risk review under CAA section 112(f): that interpretation would allow the EPA to avoid the cost/benefit analysis that is required when revising standards under section 112(d)'s recurring technology review.

"Cost is almost always a relevant—and usually, a highly important—factor in regulation. Unless Congress provides otherwise, an agency acts unreasonably in establishing 'a standard-setting process that ignore[s] economic considerations.'" *Michigan v. EPA*, 576 U.S. 743, 769 (2015) (Kagan, J., dissenting) (quoting *Industrial Union Dep't v. Am. Petroleum Institute*, 448 U.S. 607, 670 (1980) (Powell, J., concurring in part and concurring in judgment)).

The residual risk review provision of section 112(f)(2) does not explicitly require the consideration of cost, but the technology review provision of section 112(d) *does* require EPA to "tak[e] into consideration the cost of achieving such emission reduction[.]" 42 U.S.C. § 7412(d)(2). An implied discretionary authority to conduct recurring risk reviews would thus allow EPA to circumvent the cost considerations that Congress deliberately built into the technology review process, opening a significant loophole in the statutory scheme.

Section 112's explicit, carefully crafted cost/benefit framework also answers a core concern that EPA and environmental groups have previously raised: whether EPA would be able to address newly discovered or changed risks. The technology review process under section 112(d)(6) enables the EPA to consider whether further emission reductions are "necessary"—a broad term that allows the agency to consider issues related to public health and risk mitigation—while still ensuring that revisions of standards under that provision account for whether additional or improved emissions control technology is available to the industry, and its associated costs (not only to acquire that technology, but to install, maintain, and operate it). *See* 42 U.S.C. § 7412(d)(6). The EPA's authority to revise standards under section 112(d) thus allows the Agency to address new or changed risks, but without disregarding the practical considerations (including but not limited to costs) that, as explained above, made section 112 workable in the first place.

That, in turn, prevents the EPA from setting standards that violate the D.C. Circuit's admonition when addressing CAA section 112 in *United States Sugar Corp. v. EPA*, 113 F.4th 984, 994 (D.C. Cir. 2024) (internal quotation marks and citations omitted), that "we should not lightly assume that a statute is draconian . . . or demands the impossible." The EPA's interpretation of section 112 in the 2024 Rule had precisely that effect; it caused the Agency to impose draconian, impossible-to-meet standards on sterilizers under section 112(f), as the President recognized when he issued exemptions to avoid a breakdown in the supply chain for medical devices and other critical sterilized products.

EOSA suggests that the EPA include this additional reason why its revised interpretation of section 112(f)(2) better comports with Congress's intent, the statutory scheme, and common sense in its Final Reconsideration Rule.

5. CAA Section 307(d)(1)(C)

EOSA agrees with the EPA that section 307(d)(1)(C) of the CAA does not give the Agency open-ended authority to revise risk-based standards set under section 112(f)(2). *See* 91 Fed. Reg. at 12,713. But EOSA respectfully submits that the EPA's reason for so concluding does not fully comport with the text of that provision, which does not give the EPA authority to revise standards set under section 112(f) at all.

Section 307(d)(1)(C) identifies CAA provisions to which the rulemaking requirements set forth in subsection (d) apply:

> [i] the promulgation or revision of any standard of performance under section 7411 of this title, or emission standard or limitation under section 7412(d) of this title, [ii] any standard under section 7412(f) of this title, or [iii] any regulation under

> section 7412(g)(1)(D) and (F) of this title, or any regulation under section 7412(m) or (n) of this title[.]

42 U.S.C. § 7607(d)(1)(C).

Notably, the phrase "promulgation or revision"—on which the EPA relied in the 2024 Final Rule to suggest it has broad authority to do seriatim risk reviews under section 112(f)(2)—is used only once, at the front of the provision. That phrase is then followed by a list of CAA standard-setting provisions that includes, in the middle, section 112(fb):

1. "the promulgation or revision of any standard . . . under section 7411 . . . or . . . section 7412(d)";
2. "any standard under section 7412(f)"; and
3. "any regulation under section 7412(g)(1)(D) and (F) . . . or . . . under section 7412(m) or (n)[.]" (*Id.*)

It would be unusual and contrary to the last-antecedent canon—the "ordinar[y]" rule when interpreting a series of phrases—to read the phrase "promulgation or revision," which is used immediately before only the first of those items, as applying to all three. *See Lockhart v. United States*, 577 U.S. 347, 351 (2016) (the "'rule of the last antecedent' . . . provides that 'a limiting clause or phrase ... should ordinarily be read as modifying only the noun or phrase that it immediately follows'"); *United States v. Pritchett*, 470 F.2d 455, 459 & n.9 (D.C. Cir. 1972) (the last antecedent rule is "one of the simplest canons of statutory construction" and the "normal reading of the language of the statute" unless context indicates otherwise). EOSA therefore does not agree with the EPA's reading of the paragraph as referring to "the promulgation or revision of . . . any standard under section [1124(f)(2)][.]" 91 Fed. Reg. at 12,713 (quoting 42 U.S.C. § 7607(d)(1)(C)).

However, the EPA reaches the correct result because, even if the phrase at the outset of subsection (d)(1)(c) applies to the whole, it would not follow that every type of CAA standard listed thereafter can be revised at EPA's whim. Rather, some of the listed standards might only be "promulgat[ed]," while others might be "revis[ed]." 42 U.S.C. § 7607(d)(1)(C).

Further, EOSA agrees with the EPA that, even if subsection (d)(1)(c) may apply where the EPA is revisiting a standard set under section 112(f) pursuant to a court remand (i.e., after risk-based standards were found unlawful or insufficiently supported upon judicial review), that does not mean that the EPA has general authority to undertake multiple rounds of risk review and set or revise risk-based standards under section 112(f) at any time. The rulemaking process conducted upon such a remand (which often would be coupled with vacatur of the unlawful standards) would relate back to EPA's initial and sole round of risk-review, and thus the EPA is more correctly viewed as promulgating risk-based standards in the first instance in that situation.

Ultimately, the scope of the EPA's authority under any of the listed provisions, including section 112(f), must be determined based on the text and context of that particular provision. And "Congress, [the Supreme Court has] held, does not alter the fundamental details of a regulatory scheme in vague terms or ancillary provisions—it does not, one might say, hide elephants in mouseholes." *Whitman v. Am. Trucking Associations*, 531 U.S. 457, 468 (2001). Section

307(d)(1)(C), a comparative "mousehole" that simply lists other sections of the Act to which general rulemaking procedures apply, does not itself enlarge the EPA's authority by permitting what section 112 itself does not: recurring rounds of risk review.

6. CAA Section 301

EOSA anticipates that some commenters may argue that the EPA has general authority under CAA section 301 to revise its own regulatory standards when it deems necessary, including those set under section 112(f)(2). *See* 42 U.S.C. § 7601(a)(1) (providing general authority "to prescribe such regulations as are necessary to carry out [its] functions"). EOSA submits that section 301—another general and generic provision like section 307—cannot rightly be read as granting the EPA authority to revise emission standards where the substantive CAA provision describing the process and requirements for those particular standards does not.

In 2006, when the EPA conducted its first technology review and one-time residual risk review, it claimed a continuing "authority to revisit (and revise, if necessary) any rulemaking" under CAA section 301. *See* 71 Fed. Reg. 17,712, 17,715 (Apr. 7, 2006). The Proposed Reconsideration Rule explains that this statement is better understood as a reference to EPA's authority under section 112(d)(6). 91 Fed. Reg. at 12,713. EOSA believes that the EPA can better explain why section 301 does not give the EPA open-ended authority to undertake a second (or any subsequent) round of risk review under section 112(f) by identifying cases in which the D.C. Circuit rejected past attempts to rely on that generic provision for carte-blanche authority.

For example, in *Natural Resources Defense Council, Inc. v. Reilly*, 976 F.2d 36 (D.C. Cir. 1992), where the EPA had established emissions standards for radionuclides under CAA section 112(b), but then stayed implementation of the final standards for three months—and then further stayed the standards another three times—the EPA claimed its authority to issue such stays derived from its section 301 general rulemaking authority. *Id.* at 38-39. The D.C. Circuit soundly rejected that argument: "[T]he [CAA] mandated a highly circumscribed schedule for the promulgation of regulations" and, "[i]n the face of such a clear statutory command, we cannot conclude that section 301 provided the EPA with the authority to stay regulations that were subject to the deadlines established by section 112(b)." *Id.* at 41. Other decisions addressing section 301 are in accord. *See, e.g.*, *Nat. Res. Def. Council v. EPA*, 749 F.3d 1055, 1063 (D.C. Cir. 2014) ("EPA's [section 301] authority to issue ancillary regulations is not open-ended, particularly when there is statutory language on point"); *Citizens to Save Spencer Cnty. v. EPA*, 600 F.2d 844, 873 (D.C. Cir. 1979) (CAA section 301 "does not provide the Administrator with Carte blanche authority to promulgate any rules, on any matter relating to the Clean Air Act, in any manner that the Administrator wishes").

Here, too, a detailed statutory scheme specifies how, when, and in what order the EPA is to regulate under section 112. The EPA should explain in its Final Reconsideration Rule that, in accordance with the caselaw applying section 301 and circumscribing EPA's authority thereunder, section 301 cannot be employed to allow the EPA to conduct additional rounds of risk review and set or revise new section 112(f)(2) standards. That deployment of this general rulemaking provision would be contrary to the Act's statutory scheme.

Questions 3 through 6: Overarching legal issues with the EPA's reliance on the 2016 IRIS.

In response to the EPA's questions 3, 4, 5, and 6, which address EtO risk assessment for purposes of regulation under section 112, we provide the below comments addressing cross-cutting legal issues. The technical aspects of these questions are addressed in Volume 3 of EOSA's comments.

EOSA agrees that the EPA should not rely solely on the 2016 IRIS risk value in establishing regulatory standards for sterilizers emitting EtO. As the EPA recognized in its April 27, 2026 Memorandum entitled *Future Development and Use of Risk Assessments*,[1] EPA's historic approach to assessing risk has been controversial, and EPA's deployment of IRIS risk assessments as a basis for regulating under the CAA and other programs is "one of the areas of largest stakeholder concern and disagreement." Apr. 27, 2026 Memo at 3. The EPA also rightly admitted in its April 27, 2026 Memo that the EtO IRIS, and the EPA's reliance on that IRIS in rulemaking, is particularly problematic. *Id.* at 4.

Among other issues that the EPA should now address, there is a substantial body of additional studies and data that must be considered when assessing EtO risk for purposes of setting regulatory standards, as discussed in Volume 3 of EOSA's comments. Here, EOSA addresses cross-cutting legal issues raised by the EPA's reliance on the IRIS, including that (1) it was error for the EPA to rely solely on the IRIS in promulgating the 2024 Final Rule; (2) certain cross-cutting issues with the EtO IRIS render the EPA's reliance on it arbitrary and capricious; and (3) it is problematic for the EPA to continue to rely on the 2016 IRIS to determine the cost-effectiveness of technology-based EtO emission standards under section 112(d).

1. Congress was clear that the EPA must consider all relevant data and studies when regulating under section 112.

The legislative history and the EPA's interpretations of the 1990 CAA Amendments make clear that Congress prohibited the EPA from relying solely on an IRIS. Rather, Congress required that the EPA consider a range of values and incorporate uncertainties in determining a risk value for use in Section 112 rulemakings. The question of how to assess risk thus is not only a technical determination, but an issue of statutory interpretation.

Given Congress's incorporation of the Benzene NESHAP into section 112(f), that rule informs issues of statutory interpretation.[2] There, the EPA considered but rejected EPA's reliance only on the "maximum individual risk." Rather, the EPA noted that in making regulatory determinations, it would need to consider "the science and policy assumptions and estimation uncertainties associated with the risk measures," as well as the weight of the scientific evidence for human health effects." *Id*. at 38,046. Notably, the benzene NESHAP did not rely upon an IRIS value.

Frustrated with the EPA's slow pace of setting risk-based standards, Congress amended section 112 in 1990. That revised section required that the Administrator report to Congress on:

[1] From D. Fatouhi to EPA General Counsel, Assistant Administrators, et al. (Apr. 27, 2026).
[2] 42 U.S.C. § 7412(f)(2)(B) (referencing the benzene NESHAP).

(1) methods for calculating remaining risk after imposition of MACT controls; (2) "public health significance of remaining risk" and methods and costs for "reducing such risk;" (3) information on the "actual health effects" and any available epidemiological or other health studies, risks presented by background concentrations of hazardous air pollutants, [including] any uncertainties in risk assessment methodology or other health assessment technique"; and (4) legislative recommendations on addressing such risks.[3] If Congress did not act on those recommendations, the EPA Administrator was instructed to promulgate standards following the benzene NESHAP. Accordingly, the 1990 Amendments adopted an approach mandating that the EPA may not rely on a single value, but rather must account for all uncertainties and other information relevant to assessment of risk.

The EPA's contemporaneous and near-contemporaneous statements comport with this view of the Agency's obligations. First, in the report mandated by the statute, the EPA told Congress that it would "not be relying exclusively on IRIS values" but instead would "be considering all credible and readily available assessments." Moreover, the EPA noted: "In more refined assessments, which may become the basis for a risk management regulatory decision, we will consider all credible and relevant toxicity information."[4] The EPA thus made clear to Congress that, as Congress had instructed, it would be considering more than IRIS when setting standards under section 112.

Moreover, even before that submittal, the EPA had already made clear that IRIS values could not be solely relied upon for Section 112 implementation generally. An August 1994 memo from the Director of the EPA's Office of Air Quality and Planning Standards to staff on implementing early reduction credits under 112(i)(5) noted:

> It is also important to remember the IRIS is not a comprehensive toxicological data base. There may be more recent, credible and relevant information available than is contained in IRIS. Moreover, the act of including a value in IRIS is not subject to notice and comment rulemaking, and may not necessarily have been subjected to external peer review (although some specific values are based on information developed through a notice and comment rulemaking in a particular context).
>
> Accordingly, IRIS values are not entitled to conclusive weight and shall not be made legally binding in the context of any other rulemaking action. In addition, EPA or any State agency that uses IRIS should not rely exclusively on IRIS values but should consider all credible and relevant information that is submitted in any particular rulemaking. If an outside party questions IRIS values during the course of an EPA proceeding (for example, during the evaluation of a hazardous air

[3] 42 U.S.C. § 7412(f)(1).

[4] EPA, Residual Risk Report to Congress, at 57 (Mar. 1999), *available at* https://www.epa.gov/sites/default/files/2013-08/documents/risk_rep.pdf (last visited Apr. 16, 2026).

pollutant), EPA will consider all credible and relevant information before it in that proceeding.[5]

The EPA continued to issue similar warnings both in the context of Section 112 and elsewhere, thus recognizing what the statute and its history make clear: the EPA's obligation to consider all available information and account for all uncertainties brought to its attention is not a policy choice, but rather a statutory obligation embedded within section 112.

2. There are overarching issues with the 2016 IRIS that render reliance on it for purposes of setting emission standards arbitrary and capricious.

There are several other overarching issues with the EPA's reliance on the 2016 IRIS in the face of what the EPA recognized is very significant uncertainty associated with EtO's carcinogenicity and risk profile. Those issues, including that ambient concentrations of EtO are well above the risk threshold selected by the EPA and that the EPA wrongly assumed that EtO is antagonistic rather than additive if mixed with other chemicals, render any reliance on the IRIS arbitrary and capricious.

In promulgating the Miscellaneous Organics NESHAP (MON) Rule, the EPA "acknowledged that uncertainties regarding the magnitude of EtO's carcinogenic potential could materially affect risk estimates."[6] The EPA included a technical memo in the MON docket that identified two sources of uncertainty: model selection and use of upper confidence intervals (CIs) versus the central estimate. The model selected is conservative, as is the use of upper CIs, thus compounding conservatisms. Based on that technical memo, the EPA concluded that the EtO IRIS value could overstate risk by as much as a factor of five, and thus determined that risks from the MON source category were acceptable not withstanding that calculated risks were above 1-in-a 10,000 (the presumptive threshold for risk acceptability).[7] When the EPA issued the final sterilizer rule in 2024, however, it ignored this up-to-5-fold overstatement of risk when determining the level of control to impose with no explanation for this change in position.

The EPA now seeks comment on the above previous approach, as well as any new information regarding EtO's carcinogenic potential. The Agency's consideration of uncertainties in the 2024 Rule was overly constrained by several factors, including: the weight it gave to the IRIS value, failure to consider uncertainties presented by commenters, failure to engage with information on biological plausibility (which provides the basis for causal inference), and by its decision to ignore the CHP model in trying to determine what value to apply for risk management purposes. We provide detailed explanations of these failures in Volume 2, but discuss here two issues in particular—ambient concentrations and co-exposures—which demonstrate that the EPA's reliance on the 2016 IRIS was arbitrary and capricious, contrary to the EPA's obligation to engage in "reasoned decisionmaking." *See Motor Veh. Mfrs. Ass'n v. State*

---

[5] John Seitz, Director of EPA's Office of Air Quality and Planning Standards, *Guidance on the Use of Integrated Risk Information System (IRIS) Values* (Aug. 26, 1994) ("IRIS values are "only a starting point for risk assessment" and "not meant to replace careful thought and analysis necessary" for making regulatory decisions.").

[6] 91 Fed. Reg. at 12,714.

[7] 85 Fed Reg, 49,084, 49,096 (Aug. 12, 2020) (finding risk of 200 in-a-million acceptable).

Farm Ins., 463 U.S. 29, 52 (1983); *see also Dep't of Homeland Sec. v. Regents of the Univ. of Cal.*, 140 S. Ct. 1891, 1912–13 (2020).

i. The EPA's failure to account for ambient concentrations

Throughout the IRIS development process and in promulgating various regulatory actions, commenters have raised with the EPA that ambient concentrations of EO are orders of magnitude above the IRIS value. This issue has been raised both as a reality check during the IRIS development process, as well as subsequently in the EPA's use of that IRIS value for regulatory purposes.

The EPA's consistent response has been that it is regulating risk above background and therefore ambient concentrations are irrelevant. This view, that ambient concentrations are not relevant in determining the reasonableness of using a value for regulatory purposes cannot be reconciled with what the EPA told Congress. The Residual Risk Report contains an entire section (4.2.2) on addressing Risks Posed by Background Concentrations. That section discussed how other parts of the EPA address background concentrations, and noted that: "If the relative contribution of background risk is high compared to the incremental residual risk, additional source risk reduction may provide relatively negligible benefit."[8] Indeed, the EPA told Congress that in such circumstances risk from background concentrations might directly influence control requirements.[9]

The EPA's steadfast refusal to address ambient concentrations, as well as its explanation for why it refuses to do so (that the EPA is addressing excess risk) is simply inconsistent with its explanation to Congress. The refusal to address background in the case of EtO is particularly egregious since, in addition to ambient EtO in the environment, the human body produces EtO internally (endogenous EtO). Any reasonable risk assessment must fully account for and reconcile risk values with both background environmental and endogenous EtO concentrations.

Indeed, the EPA recently recognized the problematic nature of relying on an EtO IRIS value that is "at least 10,000 times lower than levels naturally occurring in the human body." Apr. 27, 2026 Fatouhi Memo at 4. And the EPA pointed to its admission in the proposed reconsideration rule that is the subject of these comments that addressing "two sources of uncertainty in the IRIS value for ethylene oxide alone … resulted in an EtO IRIS value that was five times lower (i.e., EtO is five times safer than the IRIS value provided)." *Id*. Given these admissions, EPA cannot reasonably continue to rely on the 2016 IRIS to set regulatory standards for EtO emissions.

ii. The EPA's flawed approach to assessing risk from co-exposures

In the Residual Risk Report to Congress, the EPA explained that the general default assumption for carcinogens is that they are additive. This was not surprising, at is consistent with

[8] Residual Risk Report of Congress at 94.

[9] *See id.* ("In such cases, the relative contribution of background to the total risk from HAPs would be considered in decisions for more stringent regulation and may influence the level of reductions required to obtain an "ample margin of safety.") It is clear that EPA is referencing the entire residual risk process (since the is how Congress references it in the statute), not simply the second step of the regulatory process.

the EPA's own 1986 Guidelines for the Health Risk Assessment of Chemical Mixtures which state: "If sufficient data are not available on the effects of the chemical mixture of concern or a reasonably similar mixture, the proposed approach is to assume additivity."[10]

Despite this default and its own Guidance, the EPA rejected consideration of the DOW Union Carbide study, as well as information demonstrating a lack of lymphoid cancer in smokers based on co-exposures. The EPA did so without any evidence supporting an antagonistic effect. Rather, the EPA presumed that, in certain circumstances, *antagonistic effects* may occur—without providing any evidence to support that assumption. Not only does this unsupported assumption confirm the uncertainty inherent in the EPA's assessment, but it also runs counter to the default assumption that the EPA indicated applies for carcinogens and explicitly states for chemical mixtures: risks are to be treated as additive, not antagonistic, in the absence of contrary evidence.

Indeed, the EPA explicitly noted to Congress that, for carcinogens with a linear mode of action, it would generally *add* the risks associated with them in determining if actions were needed to provide an ample margin of safety.[11] Further, if other chemicals are antagonistic to EtO-induced cancer, this logically raises an additional question regarding the relevance of the NIOSH study to any EtO exposure where there are any meaningful co-exposures, thus limiting the applicability of the dose-response and resultant IRIS cancer value to no cohort other than the NIOSH cohort. This is yet another reason why it was arbitrary and capricious for the EPA to rely on the IRIS in setting standards for sterilizers under section 112—a mistake the EPA should not repeat in this reconsideration rulemaking.

3. The EPA should not rely on the 2016 IRIS when assessing the cost-effectiveness of technology-based standards.

The EPA is right to question the EtO IRIS and recognize the significant uncertainty associated with EtO's carcinogenic potential. But EPA should also recognize that, from a cost/benefit perspective, the technology-based controls that the EPA proposes to impose are inconsistent with that aspect of the proposal.

The EPA requests comment on its proposed determination that "significant uncertainties regarding the magnitude of EtO's carcinogenic potency, particularly at low concentrations, would be an additional reason for rescinding the EtO standards."[12] As discussed above, we agree with and support that determination. That determination, however, has direct implications not only for whether EPA should promulgate standards under 112(f) (which, as discussed above, it lacks statutory authority to do in any event), but also the Agency's analysis of what technology-based standards are appropriate under 112(d). The EPA should recognize that there is a

[10] *Guidelines for the Health Risk Assessment of Chemical Mixtures*, EPA/630/R-98/002 (September 1986) at § 3.2 (p.14, first sentence).
[11] Residual Report to Congress at 124 ("In assessing cancer risk in the refined analysis, multiple HAP exposures are treated as additive where scientifically appropriate or in the absence of information to the contrary (consistent with EPA policy), and the numbers of people exposed in various subpopulation groups may be considered.")
[12] 91 Fed. Reg. at 12,714.

disconnect between its proposal to revisit the IRIS and its continued treatment of EtO as a "highly toxic HAP" in its 112(d) analysis.

In promulgating standards under 112(d), the EPA has often, as a way of considering cost, looked at cost effectiveness, "which is the relationship of costs to emission reductions, because that indicator sheds the most light on whether the revised emissions standard that is based on those controls is 'necessary' under CAA section 112(d)(6)."[13] In the 2024 Final Rule, the EPA imposed controls that, on average, had a cost effectiveness of $3.98 million per ton of EtO emitted.[14] This reflected the EPA's then-view of EtO as highly toxic, which in turn was based on the 2016 IRIS, which as discussed above suffers from fundamental legal flaws, in addition to failing to reflect all relevant data and studies (as discussed in Volume 2 of these comments).

Now, while proposing that the IRIS value is not appropriate for regulatory purposes, the EPA continues to apply a cost effectiveness metric, $2.53 million per ton on average, that aligns with its former view of EtO as a "highly toxic HAP" rather than a typical organic HAP. We have concerns about the EPA's approach to "highly toxic HAP" generally, but in any event EPA's implied determination that requiring emissions reductions with such associated costs is reasonable is inconsistent with its proposed determination that reliance on the extant IRIS is improper. We suggest that the EPA recognize that disconnect even if the EPA proceeds to finalize the technology-based standards as proposed, and that the EPA may need to take further action in the future to revise standards based on a revised IRIS value and associated lower cost-effectiveness metric.

[13] 91 Fed. Reg. at 9,096.
[14] $83.6 million annualized cost/21 tons of EtO emissions reduction.

Volume 2 (A)

Reader's Guide to Comments on the EtO risk value

The following pages are intended to guide EPA decision-makers as they review scientific comments on how cancer risks associated with EtO should be considered in this rulemaking. This guide does not address every issue raised in the detailed comments that follow. Instead, it is meant to help readers understand that the IRIS model is fundamentally flawed and unreliable, and that scientific research and analysis developed since its finalization support the use of far more plausible alternative models.

As described in EOSA's comments, these alternatives include models based on animal data and models grounded in the NIOSH study. By avoiding the implausible assumptions underlying the IRIS assessment, these biologically plausible alternatives ensure robust, scientifically credible risk values that are protective of cancer risks based on the presumed genotoxic mode of action (MOA).

Crucially, the overprotection built into the current IRIS value offers zero additional public health benefit compared to these alternatives. Continued reliance on the IRIS model approach will inexorably result in realistically unobtainable regulatory targets and scientifically indefensible risk management decisions. In the case of the medical sterilization industry, this regulatory misdirection is ultimately counterproductive to public health, as it threatens to constrain the supply of sterile medical equipment needed to protect patient health.

Ultimately, we hope the reader will also understand that the EtO IRIS value, and its use, fall well short of the guideposts for Gold Standard Science recently announced by EPA's Deputy Administrator.[1]

<u>Failure to adequately consider uncertainties</u>:

The proposed rule appropriately requests any alternative values. This approach is more appropriate and consistent with the statutory framework than simply relying on the IRIS value. While an important development, EPA's recognition that risk may be overstated by a factor of 5, is not based on consideration of the full universe of information available necessary to quantify the uncertainty. For example:

(1) EPA continues to consider the 2-piece linear spline model Cox Proportional Hazards (CPH) (2-piece spline or IRIS model) as scientifically sound and thus uses it as the starting point for its uncertainty analysis.

(2) EPA does not consider standard log-linear CPH model (TCEQ's model approach) CPH, despite having similar statistics and being more biologically plausible.

(3) EPA did not consider EPA IRIS cancer risk values derived from animal data, in which exposures are controlled. These animal data could further inform (if not replace) the NIOSH data, which suffers from exposure uncertainties; and.

(4) EPA did not account for uncertainties raised in the comments about the inaccuracies in EPA's justification for its IRIS value. [2]

---

[1] From D. Fatouhi to EPA General Counsel, Future Developments and Use of Risk Assessments, Assistant Administrators, et al. (Apr. 27, 2026). ("IRIS memo")

[2] Considering this full range of information is also consistent with the IRIS memo

Together, these omissions indicate that the EPA IRIS narrow focus results in an inaccurate characterization and estimation of uncertainty.

Significant newly developed information since 2016 supports a shallow, rather than steep, slope:

Newly developed data and analysis since 2016 further supports that the appropriate cancer slope for EO is shallow, rather than steep. This includes additional analysis of studies on endogenous production of EtO, additional ambient concentration data demonstrating background EtO levels orders of magnitude above the IRIS value, and an analysis of EtO concentrations in cigarette smoke and a comparison against observed lymphoid cancer rates in smokers. These new lines of evidence are precisely the type of readily available information that the Deputy Administrator noted "that might affect the dose-response values" that IRIS had refused to consider.[3]

(1) Endogenous: As the Deputy Administrator recently noted, EtO is produced in the body at orders of magnitude higher than the IRIS value.[4] Numerous studies demonstrate that the body produces EtO at levels tens of thousands of times the EtO IRIS value. If the IRIS risk is assumed as correct, changes in EtO exposures in the IRIS value range are not meaningfully differentiated from total background exposures naturally produced in the human body and thus do not plausibly contribute to additive cancer risk above background. A recent paper suggesting there is substantial uncertainty in sources (endogenous and ambient air) contributing to total background EtO exposure reaches its conclusion by relying on a less reliable analytical method and excluding a significant number of studies inconsistent with their position.

(2) Background Ambient concentrations: EPA's own background air measurement data, including that collected using the updated test method continues to show ambient concentrations orders of magnitude higher than the IRIS value and well above the MDL. Other than asserting a lack of confidence in it, EPA has not explained why that lack of confidence is so high as to render inappropriate the implications of that data—that ambient EtO concentrations are orders of magnitude higher than the EtO IRIS value. This has direct implications both for model selection (is the 2-piece spline model biologically plausible) and from a standard setting standpoint. As EPA wrote in the residual risk report to Congress, while IRIS values measure additional risk, "If the relative contribution of background risk is high compared to the incremental residual risk, additional source risk reduction may provide relatively negligible benefit."[5] Recent data confirms this to be the case near sterilizer facilities and this lack of benefits is something EPA has never addressed, but is obligated to address in setting standards.

(3) Biological Plausibility: The SAB instructed EPA to ensure that any model selected is biologically plausible, which animal data can help inform. Recent new animal studies evaluating key EtO mode of action events (systemic exposures, DNA adducts, apical mutagenicity) over a 4000-fold range of exposures demonstrates EtO conservatively exhibits a single linear and shallow dose response consistent with a preferred standard log-linear CPH risk model, and not the IRIS two-piece spline model.

---

[3] IRIS memo at 4.
[4] IRIS Memo at 4 (noting that it is at least 10,000 times higher).
[5] Residual Risk Report of Congress at 94.

Other human and animal studies likewise support the standard CPH approach as they show a shallow initial slope. EPA IRIS (2016) never considered the EtO mode of action or other EtO toxicology data when considering biological plausibility of statistical models to apply to the NIOSH data[6]:

a. EPA's main rationale for biological plausibility is that the EPA's chosen model is less steep than other models with excessively steep initial slopes. However, contrasting with models already deemed biologically implausible does not establish biological plausibility for EPA's selected model.

b. EPA IRIS did not provide any explanation for why the standard log-linear CPH model lacks biological plausibility for a direct-acting mutagen. In fact, the log linear CPH model approach is far more consistent with the EtO-specific MOA data than the EPA's selected model.

c. The fact that some chemicals show a steep initial slope that later flattens due to saturation of a metabolic activation pathway also does not support applying a similar spline-model shape to EtO, which does not share that mechanism.

(4) Smokers: Cigarette smoke contains analytically measurable EtO and has been used in a new study to validate that EtO exposure, as measured by its specific hydroxyethylvaline-hemoglobin adduct, exhibits a continuously linear dose-response between very low ppb exposures contributed by smoking and very high ppm-level occupational exposures. If the IRIS value is assumed as a reliable indicator of additive cancer risk, it predicts a statistically significant increase in LH would be detected in smokers. Detailed analysis based on The Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial does not show a statistically significant increase in LH cancer in any of the groups of smokers (light to heavy).

The regulatory history of the IRIS value is rife with examples of the Agency dismissing information that "might affect dose response values":[7]

Both while developing the IRIS value and in subsequent response to comments, EPA IRIS has dismissed information that would undermine the IRIS value. We cover a few of these here:

(1) Exposures in the NIOSH study are not validated: Contrary to repeated assertions by the IRIS office, the NIOSH exposure model is not a validated model for the majority of exposures prior to 1978. All of EPA's cancer slope estimates cited in EPA's memo rely on the NIOSH exposure model. Hornung et al. (1994), the developers of the NIOSH exposure regression model, explicitly stated that they had little exposure data from 1940-1978. As discussed by EOSA comments, for lymphoid cancer over 70% of the total exposures, and almost 90% of exposures below the knot occurred prior to 1978. The exposure model was well validated for exposures after 1978, but Hornung et al. (1994) altered the model for estimating exposures prior to 1978 by holding a key variable constant to 1978 levels. Hornung et al. (1994) assumed incorrectly that there were no changes in industrial hygiene practices, worker-safety procedures or leakiness of equipment prior to 1978.. EPA's position that the NIOSH exposure regression model is well-validated is just that, an

[6] The deputy Administrator explicitly called out the IRIS program for its failure to consider MOA information. IRIS Memo at 4.
[7] Id. at 4.

assumption, and an erroneous one at that. Accordingly, the claim of validation is in error for the vast majority of exposures. That the model could be validated post 1978 simply does not validate pre-1978 exposures.

Moreover, the material submitted to EPA in the docket for multiple rulemakings demonstrates that practices changed significantly in ways that would likely have impacted the exposures regardless of job description. For example, air wash cycles prior to removing sterilized material from sterilizer chambers increased over time reducing off gassing into areas where employees would be. This would reasonably be anticipated to significantly reduce exposure over time, which would reasonably be anticipated to change the model. While the submitted material may not have been sufficient to derive another exposure model, it conclusively demonstrates that the NIOSH exposure model is not a high quality validated model because it fails to account for significant process changes prior to 1978. The implications:

a. EPA needs to give serious consideration as to whether the NIOSH study can be used for quantitative and qualitative risk evaluation;

b. If it decides to continue to use it to inform risk analysis, EPA must:

   i. recognize that the DOW/UC study likely has better exposure data and incorporate that study into model selection considerations;

   ii. Recognize that it cannot select a steep slope in support of higher exposure risk given the lack of validation since lack of high-quality validated models can falsely suggest a steep slope. As EPA's cancer guidelines caution "a steep slope indicates that errors in an exposure assessment can lead to large errors in estimating risk."

   iii. Account for the likely understatement of exposure in any risk management decision.

(2) Breast Cancer:

a. IRIS has relied in part on TCEQ's decision to not include breast cancer as a basis for rejecting TCEQ's CPH model for lymphoid cancer. They are independent issues. EPA could, for example, conclude that the standard log-linear CPH model applied by TCEQ to lymphoid cancer is appropriate, reject TCEQ's decision to not account for breast cancer and then select a standard log-linear CPH model to the breast cancer data. Table 2 includes such values.

b. EPA should focus on breast cancer mortality rather than incidence. Based on information provided by Steenland et al. (2003) there are a substantial number of missing cases, likely a result of an inability to locate shorter-term workers

c. The NIOSH exposure model issues described above also applies to breast cancer mortality and breast cancer incidence data.

(3) Visual Fit: EPA has continued to rely on visual fit against the categorical result to justify model selection, claiming that that two-piece spline fits the categorical result better and rejecting the CPH model as being too low. While as described below there are numerous problems with this approach, two are worth keeping in mind while reviewing these comments:

a. The apparent agreement or disagreement between categorical estimates and continuous model predictions along the y-axis is not meaningful because each continuous model is normalized to its own model-specific baseline hazard, HR(0), estimated from the model equation using all observations. Relative rates are defined as RR(d) = HR(d)/HR(0), such that each model is scaled to its own y-intercept. Differences in vertical position therefore reflect differences in normalization rather than differences in fit to the underlying data. In contrast, categorical estimates are derived from grouped data and do not share the same model-specific baseline hazard or normalization as the continuous models. As a result, comparisons between categorical and continuous results along the y-axis are not valid. EPA IRIS (2016) explicitly recognized this limitation. However, EPA's subsequent analyses (U.S. 2022, 2024) nevertheless compare these results numerically along the y-axis, which is not appropriate.

b. EPA has also insisted that it is only looking at general shape of the curve and not location along the Y-axis. Descriptions of the CPH model as being "too low" and responses in the final rule and the HON rulemaking demonstrate that EPA is in fact numerically comparing along the Y-axis. Equally important, the general shape of 5 grouped exposures does not represent the shape of the 53 lymphoid cancers.

(4) Co exposures: EPA dismisses the DOW/UCC study and has not addressed the new EtO smoking exposure validation data, both of which do not show a clear association between EtO exposure and lymphoid cancer, let alone the high cancer rates predicted by the IRIS value. EPA speculates that both exposure scenarios include co-exposures possibly antagonistic to EtO carcinogenesis. But this is mere speculation and runs counter to EPA's default assumption that cancer risks are additive in the presence of co-exposures without a clear mechanistic explanation as to why they would be antagonistic. In the face of this information EPA should have been questioning the validity of the IRIS value, and not theorizing without evidence that the co-exposures are anti-carcinogenic.

(5) CPH model does not explode up: EPA recently indicated that it was dismissing the CPH model because it "explodes upward" at higher exposures. The CPH model, however, is essentially linear throughout the range of cumulative exposures measured by the NIOSH study for individual cases.

The above key scientific issues relevant to EPA's questions 3-6 are highlighted to aid EPA as it reviews the comprehensive scientific analysis of flaws in the 2016 IRIS and new data and studies relevant to EPA's assessment of cancer risks that follows in Volume 2(b).

Volume 2 (B)

USEPA Proposed Reconsideration of the NESHAP for Ethylene Oxide Emission Standards for Sterilization Facilities Residual Risk and Technology Review.

Response to USEPA's Request for Any Alternative Values More Appropriate and Consistent With the Statutory Framework

Response to USEPA Questions 3, 4, 5, 6

Authors:

Abby Li Ph.D., James Bus Ph.D., DABT, Fellow ATS, Kenneth Bogen, Dr.P.H., DABT, and Rick Reiss Sc.D., Exponent Inc.

B. Bhaskar Gollapudi, Ph.D., Independent Consultant

Christopher R. Kirman, M.S., SciPinion

Ciriaco Valdez-Flores, Ph.D., Independent Consultant

M. Jane Teta, Dr.P.H., M.P.H., Independent Consultant

Date:

May 15, 2026

Funding Sources:

American Chemistry Council

Ethylene Oxide Sterilization Association

EXECUTIVE SUMMARY

USEPA has solicited information on the scientific basis of the ethylene oxide (EtO) Integrated Risk Information System (IRIS) value, including advances relating occupational exposure to human cancer (Question 3), dose-response model selection issues (Question 4) and new and updated information on human exposure (Question 5). Emerging evidence, in all of these areas, shows that USEPA should not rely on the EtO unit risk value for the Proposed Commercial Sterilizer NESHAP. The IRIS unit risk value does not reflect best available science, and it is at great odds with other more credible risk estimates. A recent memo from EPA confirms that program offices can use the best available science and not rely on IRIS at their discretion (U.S. EPA, 2026).

While this document will detail a variety of scientific criticisms of USEPA's IRIS assessment, there are several things we can agree on with USEPA. It is reasonable for USEPA to assume that the putative mode of action for EtO carcinogenicity is genotoxicity, albeit weakly so in animals. The available evidence from National Institute of Occupational Safety and Health (NIOSH) studies used as basis for IRIS assessment indicates an association of lymphoid and breast cancers with high EtO exposure (Steenland et al. 2003, 2005). We also agree with USEPA's statement in the IRIS assessment that the association is "less than conclusive."

USEPA estimated a unit risk value based on a NIOSH study of sterilizer workers (Steenland et al. 2003, 2004). For cancer mortality, the original study (Steenland et al., 2004) noted "little evidence of any excess cancer mortality", "[b]reast cancer did not show any overall excess," and "[p]ositive exposure response trends for lymphoid tumours were found for males only." For breast cancer incidence, Steenland et al. (2003) concluded "Our data suggest that EtO is associated with breast cancer, but a causal interpretation is weakened due to some inconsistencies in exposure-response trends and possible biases due to non-response and incomplete cancer ascertainment." However, in 2016, USEPA released the IRIS assessment with a new reanalysis of the dose-response in the same NIOSH study based on lymphoid mortality and breast cancer incidence. The result was a 1 in a million cancer risk at a 0.1 parts per trillion (ppt) exposure, making it one of the most potent carcinogens regulated by USEPA.

Thus, the principal basis for concern about USEPA's risk estimate is the dose-response model it selected for estimating cancer risk. Our comments focus on the limitations of the NIOSH study for use for quantitative risk assessment purposes, the statistical accuracy of that dose-response modeling and the lack of consistency of the model USEPA selected with past and newly-generated biological and epidemiological evidence

USEPA has mandated Gold Standard Science which emphasizes reproducibility, transparent communication of error and uncertainty, appropriate skepticism of findings, and accepting of negative as well as positive outcomes, among other factors as key to informing environmental cancer risk assessments.

USEPA developed a cancer risk dose-response model using a 2-piece linear spline model with a knot at 1600 ppm-days cumulative exposure. This USEPA dose-response model is characterized by an initial very steep linear slope at low EtO exposures, i.e., risk rapidly increases with exposure. At the knot of 1600 ppm-days, the model abruptly transitions to a second piece exhibiting a much shallower slope at higher cumulative exposures. Other more biologically and statistically plausible dose-response models, including the more standard Cox Proportional Hazard (CPH), yield much lower potency estimates. The many problems with USEPA's dose-response modeling include:

- In calculating the significance level of its model, USEPA failed to account for the clear fact that it statistically fit the location of knot, thus reducing the degrees of the freedoms in the model by 1. When properly corrected, the model, which was only barely statistically significant to begin with, is no longer statistically significant and there is no basis for claiming that the supralinear model provides a better goodness-of-fit than the standard CPH model.
- USEPA relies on incorrect comparisons of continuous models with categorical estimates using graphs of relative rate (RR) as the primary basis for selection of the USEPA's dose-response model. USEPA compares models with the shape of the categorical estimates which do not reflect the true pattern of the individual data that are modeled. In addition, USEPA (2024) makes comparisons along the y-axis in direct contradiction to USEPA IRIS (2016) correct warning that such comparisons are invalid.
- Independent of any other issue with the IRIS analysis, the very large uncertainty associated with the NIOSH exposure assessment alone is enough to render use of the NIOSH data invalid for risk assessment. Despite the obvious flaws, USEPA IRIS barely discusses uncertainty associated with its underlying exposure estimates. Workers who worked in earlier years have a higher likelihood of being misclassified as having lower cumulative exposures than they actually received. NIOSH estimated exposures over 6 decades (1940s to 1990s) corresponding to the work tenure of the subjects, but there is no exposure data before 1976, and very little prior to 1978.

- NIOSH's exposure model implausibly predicts lower estimated exposures in earlier decades (1940s–1960s) compared with later periods (1960s–1980s) (Bogen et al. 2019). The error in the NIOSH exposure model is that NIOSH assumed that industrial hygiene and workplace standards for equipment in 1978 apply to all years previously. Yet, the 1940s to 1970s represent a period of significant improvement in industrial hygiene workplace standards with associated significant reductions in chemical exposures during this period.

- A new analysis of previously published NIOSH data[1] presented in this document shows the substantial impact of USEPA's exposure misclassification. The vast majority of the exposures for lymphoid and breast cancer mortality cases occurred prior to the period when EtO measures were available on which to validate any NIOSH estimates of occupational EtO exposures.

- The NAS report reviewing the TCEQ assessment incorrectly states that exposure errors are most likely to impact higher cumulative exposures. In fact, exposure misclassification applies to 87- 90%[2] of exposures estimated to occur below USEPA's knot for breast cancer and lymphoid mortality cases.

- The NIOSH exposure model is well-validated for exposures after 1978 but has not been validated for the majority of worker exposures between the 1940's and early 1970's. The NIOSH exposure model predicts lower exposures in earlier years compared to 1978 exposures. This is a result of the NIOSH authors altering the well-validated exposure model by holding a key variable related to industrial hygiene and worker safety constant to 1978 levels. There is ample evidence that this assumption is incorrect. As a result, workers who worked in earlier years have a high likelihood of being misclassified as having lower exposures than predicted by the model.

It is proper scientific technique to scrutinize scientific conclusions against other available evidence, and this is particularly so for USEPA's IRIS value, which yields such an implausibly high potency. We have done so using a variety of methods and data:

- There is another large cohort study – the Union Carbide Company (UCC) study of EtO manufacturing workers with an equivalent number of lymphoid cases as NIOSH. Valdez-Flores et al. (2025) concluded that this cohort does not support the steep dose-response model for lymphoid mortality predicted by IRIS.

- Integration of new Mode-of-Action (MOA) evidence characterizing DNA adduct and apical genotoxicity endpoints in mice demonstrates EtO as a weak *in viv*o genotoxicant. Pro-mutagenic O6-DNA adducts and measurable mutations occurred only at high, ppm-level exposures coincident with saturation of detoxification/repair pathways, yielding dose-response relationships that are conservatively consistent with a single-slope (log-linear or near-linear) model and provide no plausible biological support for relying on a two-piece spline with a disproportionately steep low-dose slope (Gollapudi et al., 2020, 2026; Liu et al., 2026).

[1] Valdez-Flores et al. (2010) supplementary summary tables

[2] For lymphoid mortality, there are 13 cases below the knot of 1600 ppm-days. On average 87% of the exposures for these workers were before 1978. For breast cancer mortality, there are 15 cases below the knot of 700 ppm-days. On average 90% of the exposures for these workers were before 1978.

- Although cigarette smoke contains measurable concentrations of EtO resulting in demonstrable increases in systemically absorbed EtO above natural background, the very large and extensively studied smoker cohort does not have elevated risk for lymphoid cancers. As further examined in these comments, a new analysis of a larger U.S.-based study of cigarette smokers (Troy et al., 2010) shows that the IRIS potency value, if it were true, should yield detectable lymphoid cancers that were not apparent in this smoker study.

- EtO is formed endogenously in the body from ethylene metabolism, and non-smokers have significantly higher levels of endogenous EtO, as measured by EtO-derived hemoglobin adducts. These daily naturally produced exposures are substantially higher than those contributed by EtO in ambient air. Of critical contextual importance, USEPA's risk-specific concentration is thousands of times lower than natural endogenous levels. Thus, the USEPA IRIS implausibly targets a risk-reduction range that is not differentiable from even the normal interindividual variability of endogenously generated EtO exposures. While the Agency maintains that its potency estimate applies only to "extra" risk above background, the premise that a substance is a potent carcinogen at, or immediately above, levels produced naturally by human physiology is biologically implausible. This conclusion is further supported by the additional reality check that smokers do not exhibit elevated risk for lymphoid cancer, despite having low ppb but analytically above background EtO exposures.

These analyses, using different data than NIOSH, provide a basis for critically evaluating the consistency of the USEPA IRIS value. All of these other data sources are inconsistent with IRIS.

Our review also addresses a new update of the NIOSH study alleging potentially higher breast cancer risks (Kelly-Reif et al., 2025). Kelly-Reif et al. (2025) concludes that there is a steep exposure-response at low exposures based on specific modeling choices (e.g., log–log model, 20-year lag, truncation of high exposures). Our analysis shows that these findings are questionable due to methodological issues such as potential over-adjustment for employment duration and incorrect assumption that all workers received the same lower exposures to justify truncating the data. Picciotto et al. (2026) similarly assumes without concrete evidence that there is a healthy worker survivor effect. While the authors correctly note that standard duration adjustments (like those in Kelly-Reif et al. 2025) are inappropriate, they were unable to employ the "g-methods" necessary to prove or correct the bias they allege. The updated NIOSH study suffers from the same limitations of the Steenland et al. (2003, 2004) NIOSH update in that the NIOSH exposure regression model has not been validated for the majority of exposures for this cohort between 1940-1978.

USEPA has a range of options that it can utilize to select a scientifically grounded unit risk level compared to the scientifically implausible IRIS value:

- There are several rodent carcinogenicity studies with the highest risk value being 8.33×10^{-2} per ppm from a mouse study. USEPA used this value in the Office of Pesticide Programs (OPP) risk assessment in 2008, well after the NIOSH study was published (U.S. USEPA, 2008)

- It could use NIOSH data but with more biologically plausible dose-response models (lymphoid + breast) including the log-linear CPH (5.6×10^{-2} per ppm) (as used by TCEQ), log-linear CPH (7.0×10^{-2} per ppm), or the linear CPH (0.19 per ppm).

All these alternative values are broadly consistent with one another, biologically plausible, and present EtO cancer risks substantially lower than the IRIS risk of 9.1 per ppm.

Regarding Question 6 (additional information provide in CMAS NESHAP comments), the American Chemistry Council submitted comments on the proposed CMAS NESHAP on April 14, 2026, that included an Attachment 1 regarding questions raised by ACC that have not been addressed by USEPA RTC on the MON or HON. EOSA raises these questions again to USEPA.

1. Scientific Rationale And Recommended Cancer Slope Factors For Ethylene Oxide

EOSA welcomes the opportunity to provide comments on USEPA's proposed reconsideration of the NESHAP for ethylene oxide (EtO) emissions from commercial sterilizer facilities. USEPA's Proposed Commercial Sterilizer NESHAP states that "USEPA should not rely solely on the USEPA IRIS (2016) selected risk value in establishing regulatory standards. Doing so inappropriately narrows the scientific basis for regulation and fails to reflect other more credible risk estimates available to the Agency."

There is compelling new statistical, biological and epidemiological evidence indicating that the USEPA IRIS (2016) selected risk value should **not** be included in the range of credible risk estimates. As described in our detailed response to **Question 3** below, USEPA's (2022, 2024) responses to comments (RTC) provide new evidence that the USEPA IRIS selection of the two-piece linear spline Cox proportional hazards (2-piece linear spline) model is based on incorrect visual fit comparisons, flawed statistics, and a lack of consideration of biological plausibility based on the most relevant EtO data. Furthermore, USEPA the exposure model was not validated for over 76% of exposures and almost 90% of exposures below the knot in USEPA's 2-linear spline model[3]. USEPA IRIS (2016) reports that the exposure data files supporting the NIOSH exposure regression model were lost[4]. In view of the compelling new evidence that the NIOSH exposure model underpredicts exposures in earlier years. In general, when exposures are underpredicted, cancer potency estimates are over-estimated. Thus, biological plausibility must play a primary role in determining what type of continuous Cox Proportional Hazards (CPH) model to the individual cases.

The most immediate, scientifically sound corrective action to address the high probability of exposure misclassification for the majority of cases in the NIOSH study is for the USEPA to adopt the 2016 USEPA IRIS cancer risk estimate specifically derived from animal bioassay data. At a minimum, these values should be included in the range of cancer risk estimates for EtO. The Agency should refine this analyses by utilizing USEPA's BMDS program to replace the now-obsolete ToxRisk software. The risk estimate can be further strengthened by integrating a broader weight of evidence, as proposed by Kirman et al. (2021). Combining results from the

[3] For lymphoid mortality, there are 13 cases below the knot of 1600 ppm-days. On average 87% of the exposures for these workers occurred before 1978 when there was little or no exposure data. For breast cancer mortality, there are 15 cases below the knot of 700 ppm-days. On average 90% of the exposures for these workers occurred before 1978 (Source: Valdez-Flores et al. (2010) supplementary summary tables and USEPA IRIS, 2016 Table D-34)

[4] USEPA IRIS p. H-28: In response to the panel's suggestion that the Hornung analysis represents an 'invaluable opportunity' for further analysis of the impact of possible errors in exposure estimation, Dr. Steenland visited the NIOSH offices to review the data. "Unfortunately, the electronic data files used in the exposure analysis were no longer available."

ethylene study of Hamm et al. (1984) with the EtO studies of Lynch et al. (1984) and Snellings et al. (1984**)** should be explored. The ethylene study informs the shape of the dose-response relationship for EtO tumors in rats at lower exposure than studied for EtO directly. Because ethylene is metabolized to EtO, incorporating the PBPK modeling by Filser and Klein (2018) allows for accurate dose-equivalence, showing that high-level ethylene exposures (1,000–3,000 ppm) in rats are equivalent to internal EtO doses of 5.26 and 5.52 ppm, respectively. By aggregating these datasets, the agency can perform a more robust dose-response analysis across a significantly broader range of exposure levels, reducing the uncertainty inherent in relying on any single animal study. Transitioning to this animal-based, PBPK-supported model provides scientifically superior alternatives for cancer risk values.

If USEPA is currently restricting its consideration of risk values only to those derived from the NIOSH lymphoid and breast cancer endpoints, then the standard log-linear CPH model estimates have the highest biological plausibility, with the single linear CPH model used for a biologically-credible sensitivity analysis. These model forms have greater biological plausibility and improved consistency with the broader epidemiological evidence base than USEPA IRIS selected models. The standard log-linear CPH model is well accepted by epidemiologists in cancer exposure-response analysis and is essentially linear at the observed exposures levels of the NIOSH study cases. TCEQ only applied this model to lymphoid mortality cancers, but TCEQ's model **approach** (i.e., standard log-linear CPH) has also been applied to breast cancer mortality and/or breast cancer incidence by Valdez-Flores et al. (2010), Steenland et al. (2004) and USEPA IRIS (2016). When the TCEQ approach is applied to both lymphoid and breast cancer mortality, the standard CPH model predicts overall cancer risks are approximately two orders of magnitude lower than the 2-piece linear spline model applied to both cancer mortalities.

New responses to comments (RTC) by USEPA (2022, 2024) on the MON and HON defending the USEPA IRIS use of the 2-piece linear spline model are based on incorrect statistics and visual fit comparisons on graphs of relative rate (RR). The following are examples of errors that are discussed in greater detail in our response to **Question 4**:

- USEPA (2022, 2024) explanation of USEPA's statistical methodology supporting its 2-piece linear spline model contradict and misrepresent the analysis provided by USEPA IRIS Appendix D regarding whether or not the "knot" (or bend point in USEPA's selected model) was a parameter in the model. USEPAThe knot was clearly a parameter in the model based on sequential calculation of likelihood estimates for increments of 100 ppm-days (USEPA IRIS pp. 4-16, 4-36, 4-45). The knot with the maximum likelihood estimate (MLE) was selected for breast cancer mortality and incidence data. A knot with a local MLE was selected for lymphoid mortality by searching for the next best MLE different from the true MLE.

- USEPA (2022) RTC inappropriately applies TCEQ's model (standard log-linear CPH model) to each categorical estimate separately to support their incorrect conclusion that the TCEQ model is implausible because it swings up high. Yet, this analysis has no bearing on TCEQ's actual modeling of individual estimates and misrepresents the TCEQ's model.
- The USEPA's RTCs incorrectly states that categorical and continuous models are directly comparable when expressed as relative rates. This leads USEPA to justify incorrect comparisons along the y-axis between TCEQ's model and USEPA's model despite correct warning by USEPA IRIS (2016) that such comparisons are invalid.

These and other errors in USEPA's RTCs indicate that the scientific basis for USEPA IRIS model choice is fundamentally flawed. USEPA claims they have responded to all comments, yet this is not the case as indicated in our response to **Question 6** below. Due to these errors, cancer risk values selected by USEPA IRIS (2016) based on the 2-piece linear spline model should not be used as the starting point for consideration of the range of scientifically credible risk estimates.

USEPA IRIS claims that biological plausibility was considered in selecting the USEPA IRIS final model. However, this was limited to USEPA IRIS rejection of models with initial slopes approaching infinity that are based on zero cases. USEPA IRIS did not consider biological plausibility based on EtO-specific genotoxicity or toxicology data or with respect to the standard log-linear or linear CPH model. In contrast, TCEQ (2020) grounded their model selection based on the most relevant biological genotoxicity and toxicokinetic data for EtO. New biological evidence summarized below in response to **Question 4** further supports TCEQ's modeling approach (i.e., the standard log-linear CPH model) for both lymphoid and breast cancers (Liu et al. 2026, Gollapudi et al. 2020, 2026). These new data directly contradict the steep exposure-response relationship predicted by the USEPA's model at concentrations marginally above endogenous levels. The weight of evidence based on these new findings together with existing biological and toxicokinetic data strongly supports the application of the standard log-linear CPH model for lymphoid and breast cancer endpoints.

New and existing epidemiological evidence also supports the standard log-linear CPH model. Valdez-Flores et al. (2025) provides updated epidemiological evidence that the standard CPH model more plausibly describes the relationship between EtO exposures and lymphoid mortality. This updated Union Carbide Corporation (UCC) study provides an external dataset to the NIOSH cohort with similar number of male lymphoid cases, higher average cumulative exposures, and comparable, if not superior, historical exposure data prior to 1978 than the NIOSH study. There were no statistically significant excess deaths nor exposure-response trends for any of the male cancers of interest. There were no increasing exposure-response trends based on cumulative or log-cumulative exposure metrics. Although the upper-bound slope on the non-

statistically significant CPH model was positive, it is very small in magnitude and does not indicate a steep exposure response pattern at lower exposures.

Kelly-Reif et al. (2025) is an update of the Steenland et al. (2004) NIOSH mortality study. However, the only data updated was breast cancer mortality. Lymphoid mortality, which contributes 85% to USEPA IRIS value, was not included. This update suffers from the same foundational exposure assessment limitations as the earlier Steenland et al (2003, 2004) studies. The main analysis departs from all previous NIOSH study analysis (Steenland et al. 1991, 2004) by adding an adjustment for duration of employment that has a "pronounced impact" on the model estimates. Kelly-Reif et al. claim that this adjustment is necessary to address an assumed healthy worker survivor bias (HWSB). However, Picciotto et al. (2026)**,** which shares several authors with Kelly-Reif et al., correctly states that such an adjustment for duration is an inappropriate method to address HWSB. This adjustment likely introduces overadjustment due to colinearity between the main exposure variable (ppm-days of employment) and the duration variable (years of employment).

Kelly-Reif et al. also deviates from Steenland et al. (2004) mortality study by excluding the standard log-linear CPH model (TCEQ's model approach). Instead, the linear CPH model, not previously applied by Steenland et al., was applied to the full cohort and to subsets truncated at 1000 and 5000 ppm-days. The arbitrary truncation of data sets is inappropriate for consideration for quantitative risk assessment purposes. Notably, the linear CPH model slope derived from the full cohort is lower than the linear CPH model slope estimated by USEPA IRIS (2016) for Steenland et al. (2004) breast cancer mortality. These linear model slopes are both 2-orders of magnitude lower than that of the USEPA IRIS selected 2-linear slope spline model.

Neither Kelly-Reif et al. nor Picciotto et al. acknowledge the critical lack of NIOSH worker exposure data prior to 1978. As with the Steenland et al. (2004) study, the new Kelly-Reif et al. update for breast cancer mortality relies on the NIOSH exposure regression model. As detailed in our response to **Question 3**, the majority of exposures for breast cancer mortality cases from the previous update by Steenland et al. (2004) occurred before 1978—a period for which no empirical exposure data exists[5]. Specifically, USEPA IRIS (2016) applied a 2-linear slope spline model to breast cancer mortality (using a knot of 700 ppm-days). Over 90% of exposures for cases occurred prior to 1978 for all 60 cases that were not lagged out, for the 15 cases below the knot, and for the 45 cases above the knot.

The NIOSH exposure regression model was rigorously validated for exposures after 1978; however, to estimate exposures for earlier years, NIOSH significantly altered the model by holding the **Calendar Year** variable (a surrogate for work practices) constant at 1978 levels. This "frozen" variable erroneously assumes no technological or safety improvements occurred

[5] Based on re-analyses of supplementary summary data from Valdez-Flores et al. (2010).

between 1938 and 1978. Consequently, the **Sterilizer Volume** variable became the primary driver of the equation, forcing the model to predict lower exposures in earlier years based on smaller chamber sizes—an assumption that lacks empirical validation. This assumption directly contradicts published industrial hygiene, residue level and engineering evidence (see section below on Question 3).

It is a common misconception that EtO exposure is primarily driven by industrial emissions. In fact, the U.S. population is subject to constant background exposure independent of point sources. The vast majority of this burden is **endogenous**, resulting from the metabolic oxidation of ethylene produced by gut bacteria and systemic oxidative stress. For the average individual, exogenous EtO inhalation accounts for a negligible fraction—approximately 3%—of total EtO exposure, while the remaining 97% is internally generated from metabolism of ethylene (Kirman et al. 2025[6]).

USEPA's current unit risk estimate (URE) is fundamentally disconnected from this biological reality. By setting risk-specific concentrations thousands of times lower than natural endogenous levels, the USEPA is targeting a range that is biologically insignificant and falls well within normal interindividual variability of endogenous levels. While the Agency maintains that its potency estimates apply only to "extra" risk above background, the premise that a substance is a potent carcinogen at—or immediately above—levels produced naturally by human physiology is biologically implausible. While some critics, such as Lin et al. (2025), have argued that internal production accounts for less of the background than previously thought, a subsequent critique point to flaws in their modeling, such as underpredicting known adduct levels and failing to account for diverse pathways like oxidative stress and gut microbes (Kirman and Bus, 2026). Ultimately, this body of work argues that risk management decisions must take into account the dominance of endogenous EtO and limitations of the current USEPA IRIS.

Continued reliance on the 2016 IRIS Unit Risk Estimate (URE) forces regulation of "risks" that are merely artifacts of statistical over-modeling rather than demonstrable biological hazards.

EOSA welcomes the opportunity to replace the current 2016 IRIS URE with reasonable cancer risk values.

1.1 Gold Standard Science Should Include Range of Reasonable Values

On May 2025 President Trump signed an executive order on Gold Standard science. The executive order directed the Director of the Office of Science and Technology Policy in

[6] EtO was measured as EtO - hemoglobin adducts (EtO HEV adduct), which exhibited a continuously linear relationship between low ppb smoking-related exposures up to high ppm occupational exposures, thereby addressing USEPA RTC (2022, 2024) that there are no endogenous levels.

consultation with the relevant agencies to "issue guidance for agencies on implementation of "Gold Standard Science" in the conduct and management of their respective scientific activities."[7]

Per the executive order:

> Gold Standard Science means science conducted in a manner that is:
> (i) reproducible;
> (ii) transparent;
> (iii) communicative of error and uncertainty;
> (iv) collaborative and interdisciplinary;
> (v) skeptical of its findings and assumptions;
> (vi) structured for falsifiability of hypotheses;
> (vii) subject to unbiased peer review;
> (viii) accepting of negative results as positive outcomes; and
> (ix) without conflicts of interest.[8]

Subsequently the USEPA administrator issued a memo on the implementation of the executive order.[9] The USEPA implementation document provided further detail on the definition of Gold Standard Science. As will be demonstrated in the detailed discussion that follows, the EtO IRIS value falls well short of meeting the criteria for Gold Standards Science. A few examples:

- Transparency—" Transparency in science entails the open, accessible, and comprehensive sharing of all components of the research process—methodologies, data, analytical tools, and findings—to enable stringent scrutiny, validation, and reuse by the scientific community and the public." Despite multiple attempts by third parties to get access to the full data and analysis of the NIOSH study, the critical study for USEPA IRIS quantitative risk assessment, including by use of FOIA, (or details with personally identifiable information redacted), such requests have not resulted in the access necessary to assess some of the claims made by USEPA or to verify certain counter claims.

- Communicative of Error and Uncertainty – "Communicating error and uncertainty in science entails the clear, precise, and accurate disclosure of limitations, variability, and potential sources of error or limitations in measurements or research findings, enabling other scientists to critically assess, replicate, and extend the work." This includes describing error bands on categorical grouped estimates, acknowledging that the grouped estimates are not the data modeled, and addressing Valdez-Flores and Sielken (2013) rigorous demonstration that a

[7] https://www.whitehouse.gov/presidential-actions/2025/05/restoring-gold-standard-science/
[8] *Id*.
[9] U.S. Environmental Protection Agency Implementation of Gold Standard Science (available at https://www.epa.gov/scientific-leadership/gold-standard-science

model's goodness-of-fit to the individual data should not be based on the assumption that summary estimates describe the true underlying exposure response relationship. In addition, USEPA failed to even acknowledge basic uncertainties in the key NIOSH study, including those associated with the underlying exposure model, which USEPA treats as validated not withstanding that over 70% of the data is based on unverified assumptions, and indeed ones that are refuted by the record.

- Skeptical of its Findings and Assumptions – "Maintaining constructive skepticism of findings and assumptions in science refers to the critical and open-minded evaluation of research findings, methodologies, and underlying assumptions to ensure their validity, robustness, and reliability." However, USEPA has demonstrated a distinct lack of skepticism regarding its IRIS value, as evidenced by it responses to information submitted by external parties. For example, despite unanimous agreement among TCEQ peer-reviewers that USEPA IRIS p-value calculations are incorrect, the USEPA responded with revisionist history. The agency claimed that a knot parameter in the model was "selected" based on visual fit to the lower dose data on lymphoid cancer. Yet, the official USEPA IRIS record clearly indicates that the knot position was determined through a systematic optimization approach to find the global or local maximum likelihood knot location for breast cancer incidence and lymphoid mortality, respectively. The incorrect p-values were a primary basis for eliminating from consideration the biologically plausible log-linear and linear CPH model. Because these incorrect p-values were used to exclude the more biologically plausible log-linear and linear CPH models, this example illustrates not only a failure to critically evaluate its own methodologies but also the dissemination of misleading information to the public.

- Accepting of Negative Results as Positive Outcomes – "Accepting negative results as positive outcomes in science refers to recognizing and valuing—as meaningful contributions to knowledge generation—null or unexpected findings that fail to support a hypothesis." Both the UCC study and data related to EtO exposures in smokers do not show an association with lymphoid cancer – that is, they constitute negative results that fail to support USEPA's selection of a model with an initial steep slope. Instead, these data provide positive support for the standard CPH model for lymphoid mortality cancer.

In reviewing our recommendations of alternative values and detailed responses to USEPA's Questions 3-6 below we ask that USEPA keep in mind these criteria for Gold Standard Science.

1.2. The Commercial Sterilizer NESHAP Rule Should Not Rely On The USEPA's 2019 Memorandum To Define Cancer Risk Ranges

EOSA concurs with USEPA's proposed Commercial Sterilizer Facilities NESHAP rule that "it would not be appropriate to rely on the 2016 EtO IRIS value in setting standards." The proposed rule acknowledges that the 2019 technical memorandum "demonstrates that both model selection and the underlying cancer data … can greatly influence the resulting risk estimate." However, EOSA opposes the proposed rule's reliance on the 2019 memorandum to define cancer risk ranges. The 2019 memo is insufficient for the following reasons:

- It relies solely on cancer risk estimates derived from the NIOSH study and ignores cancer risk estimates provided by USEPA IRIS (2016) based on animal carcinogenicity studies.
- It perpetuates the flawed statistical methods and misleading visual comparisons found in the USEPA IRIS analysis.
- Biological plausibility based on consideration of EtO toxicology and toxicokinetic data did not play a role in determining which models were included or excluded.
- Model selection did not consider the epidemiological evidence, specifically, the absence of effects in females and at lower doses in males in the NIOSH and UCC cohorts reported by Steenland et al. (2004, 2003) and Valdez-Flores et al. (2010).

The IRIS analysis does not apply consistent statistical criteria across models, as it fails to account for all estimated parameters in the IRIS selected model, resulting in biased model comparisons. When corrected, standard log-linear and linear Cox proportional hazards (CPH) models show goodness-of-fit comparable to the IRIS-selected 2-linear-slope spline model (Table 1), eliminating any statistical basis for favoring the USEPA's complex model over the simpler log-linear or linear models.

Table 1. Original and Corrected p-values for comparison of IRIS 2-slope linear spline and IRIS standard CPH model from the null hypothesis

	USEPA IRIS (**corrected**) 2-slope linear spline	USEPA IRIS log-linear CPH Model	USEPA IRIS Linear CPH Model
Lymphoid Mortality	**p = 0.14** corrected from 0.07 (knot at 1600 ppm-days)	p = 0.22	p = 0.13
Breast Cancer Incidence	**p = 0.04** corrected from 0.01 (knot at 5750 ppm-days)	p = 0.02	p = 0.01
Breast Cancer Mortality	**p = 0.15** corrected from 0.07 (knot at 700 ppm-days)	p = 0.10	p = 0.07

Source: Corrected and IRIS reported p-values are based on IRIS (2016; Tables 4-2, 4-4, 4-6, 4-12, 4-13, Appendix D) and TCEQ (2020). [10,11] Calculation based on USEPA IRIS Tables D-25 and D-24 values and Chi square calculation for linear model in excel =(Chisq.dist.rt(5.396,1) for 1 modeled parameter ; chi square calculation for linear spline model (Chisq.dist.rt (5.375,3) for 3 modeled parameters.

The biological evidence supports the use of the standard log-linear Cox proportional hazards (CPH) models (effectively linear over the NIOSH exposure range), rather than the IRIS two-slope linear spline model, which relies on a biologically implausible steep initial slope[12]. Despite much stronger biological grounding, the simpler standard CPH model was excluded from the 2019 technical memorandum based on flawed statistical and visual fit analysis. This lack of consideration of biological plausibility, together with underlying statistical errors, undermines the memorandum's reliability as a basis for defining the range of cancer risk assessment values for the USEPA's Proposed Commercial Sterilizer NESHAP rule.

1.3. EtO Cancer Risk Assessment Should Be Derived Primarily From Rodent Bioassay Data

The 2019 technical memorandum relies exclusively on NIOSH epidemiological studies for lymphoid cancer mortality and breast cancer incidence, omitting USEPA IRIS risk estimates derived from animal carcinogenicity data.

As detailed in our response to **Question 3**, the NIOSH studies have a high potential for exposure misclassification. While the NIOSH exposure regression model was validated for the 1978–1986 period, it lacks an empirical foundation for the 1940–1976 era due to a complete absence of historical exposure data.

Critically, the validated model was fundamentally altered for these earlier decades by fixing the "calendar year" variable at 1978 levels. This improper constraint assumes that 1940s-era industrial hygiene, air wash frequencies, equipment integrity (e.g., inferior seals and gaskets)

[10] USEPA, 2016. Evaluation of the Inhalation Carcinogenicity of Ethylene Oxide (Final Report). https://iris.epa.gov/document/&deid=329730

[11] TCEQ (2020)

[12] A more detailed explanation of the different types of CPH models are provided below in section 3.

were equivalent to 1978 standards, despite published literature to the contrary (Bogen et al., 2019). By reducing the exposure regression model effectively to a function of "sterilizer volume," it produces implausibly low exposure estimates for the 1940s–1960s that contradict the historical record. This systematic misclassification introduces broad uncertainty and a high potential for bias at the lower exposure range. Specifically, workers with high actual exposures in earlier decades are incorrectly modeled as having low exposures, making the NIOSH estimates an unreliable basis for quantitative risk assessment.

While rodent bioassays for EtO utilize fewer treatment groups and higher dose ranges, they provide **verifiable exposure levels** that offer a more reliable foundation for quantitative risk assessment. By relying on controlled, empirical measurements, these animal studies eliminate the validation gaps and back-extrapolation errors inherent in the NIOSH study.

The 2016 USEPA IRIS assessment (Table 4-20, 4-26) provides adult-based Unit Risk Estimates (UREs) ranging from **2.2×10^{-5} to 4.6×10^{-5} per µg/m³** (equivalent to **4.03×10^{-2} and 8.42×10^{-2} per ppm**). There is a high degree of concordance across species and sex. To ensure these estimates reflect the best available science, USEPA should provide updated rodent-based UREs that incorporate the following considerations:

- **Software Updates:** The ToxRisk software used in previous assessments is obsolete. USEPA should apply the latest version of its Benchmark Dose Software (BMDS) for all subsequent modeling.
- **Data Sufficiency for BMD Modeling:** In the 2016 IRIS assessment, USEPA excluded high-concentration groups from certain rodent datasets (e.g., mammary gland tumors in NTP, 1987; mononuclear cell leukemia in Lynch et al., 1984). This left only a single treatment group, which is insufficient for BMD modeling. USEPA should either exclude these datasets or retain the high-concentration groups and accept the resulting lower p-values for model fit.
- **Tumor Relevance:** USEPA should reassess the inclusion of specific tumor datasets, as the human relevance of certain endpoints—particularly mononuclear cell leukemia (MCL)—remains a subject of significant scientific debate.
- **Optimization of Model Selection and Parameters:**
    - **Dichotomous Models:** Because previous fits for the Lynch and Snellings datasets were limited to the multistage model, USEPA should evaluate additional dichotomous models available in BMDS to determine if they provide a superior fit.
    - **Polynomial Degree Flexibility:** The requirement to use a time-to-response model (multistage-Weibull) with a first-degree component should be re-evaluated. Standard multistage modeling for several tumor types (e.g., lung, lymphoma, and uterine tumors in female mice; USEPA IRIS Table G-1) required higher polynomial degrees than those currently permitted under the time-to-response model.

- **Integration of Ethylene Bioassay Data:** Dose-response modeling for rat datasets should incorporate the ethylene cancer bioassay (Hamm et al., 1984). Because ethylene is metabolically converted to EtO, internal dose levels (HEV [13]adducts or PBPK modeling) can be used to combine this study with other rat EtO bioassays to characterize the dose-response relationship across a broader exposure range. Crucially, the negative results in rats following moderate ethylene-derived EtO exposures—approximately 50–160 times higher than background—provide a critical "reality check" for USEPA IRIS dose-response assessment. These findings indicate that potency estimates based solely on high-exposure EtO-only studies significantly overestimate the response slope at low-to-moderate concentrations (see Kirman et al., 2021, Figure 7).

1.4. Unit Risk Values For NIOSH Lymphoid And Breast Cancer Mortality Should Be Based On The Standard Log-Linear CPH And Linear CPH Models

If the NIOSH study remains the critical study for cancer risk assessment, the primary cancer endpoints considered by USEPA IRIS and TCEQ for quantitative risk assessment include lymphohematopoietic (LH) cancer mortality, lymphoid cancer mortality (a subset of LH), breast cancer mortality, and breast cancer incidence from the NIOSH cohort study that was updated through 1998 (USEPA, 2016; TCEQ, 2020; Steenland et al., 2003, 2004). The overall weight of epidemiologic evidence is stronger for LH cancers than for breast cancer (Marsh et al., 2019). USEPA's IRIS (2016) cancer risk value is based on lymphoid mortality and breast cancer incidence, with lymphoid mortality contributing approximately 85% of the USEPA's risk estimate. In contrast, TCEQ (2020) relies solely on lymphoid mortality from the same cohort. We note that TCEQ's decision to not address breast cancer does not undermine its analysis with respect to lymphoid cancer. Breast cancer can be addressed using TCEQ's model approach as discussed below and as reflected in Table 2.

EOSA agrees with TCEQ (2020) that the breast cancer incidence data from Steenland et al. (2003) is not appropriate for quantitative risk assessment. However, EOSA's primary reason for excluding this data is due to the substantial under-ascertainment (missing cases) and the data are not publicly available for independent analysis. The authors of the NIOSH study indicated that workers with shorter employment duration and, hence, lower cumulative exposures were more difficult to locate. This, together with the exposure misclassifications in earlier years, indicates that this study should NOT be used for quantitative risk assessment purposes. This recommendation is consistent with USEPA Office of Pesticide Programs (OPP) Guidance[14] for use of open literature toxicity studies to support human health risk assessment. Although written by USEPA OPP, the principles are universally applicable including emphasis on adequate

[13] **N-(2-hydroxyethyl)valine (HEV) adducts** are stable biomarkers formed when ethylene oxide reacts with hemoglobin; they serve as a precise metric for cumulative internal dose

[14] USEPA https://www.epa.gov/pesticide-science-and-assessing-pesticide-risks/guidance-considering-and-using-open-literature

assessment of exposures and reasonably valid and reliable outcome ascertainment, which the Steenland et al. (2003) breast cancer incidence data lacks. In contrast, the mortality data for breast cancer and lymphoid cancers from Steenland et al. (2004) are fully ascertained, but still suffers from exposure misclassifications in earlier years. As described in greater detail below, the epidemiological, biological, and toxicological evidence support application of the standard log-linear approach to both lymphoid and breast cancer mortality data from the NIOSH cohort.

Unit risk values based on the standard log-linear CPH model were applied to the NIOSH lymphoid and breast cancer mortality (Table 2). In the majority of cases, these values could be found in TCEQ (2020) or USEPA IRIS (2016), or derived from model slopes published by Valdez-Flores et al. (2010) or USEPA IRIS (2016). As explained in greater detail below, the standard log-linear CPH model is essentially a linear model that has much greater biological plausibility based on existing and new mode of action data. The linear CPH model is included as an additional model to consider for comparison purposes. These values are derived using USEPA IRIS approach to modeling with the following exceptions:

- The point-of-departure for cancer risk assessment should be at the lower range of the NIOSH data. The POD of 1E-05 was correctly demonstrated by TCEQ (2020) to be at the lower range of the NIOSH data. In most cases, the POD of 1/100 used by USEPA is at the higher range or extrapolates above the experimental exposure range.
- Excess risks were calculated at age 70 instead of 85 years. The estimated exposure-response models are based on **worker** exposures. Calculating extra risks through age 85 years involves extrapolation of the fitted models beyond the range of the data upon which they are based. Calculating excess risks through age 85 years makes the life table analysis unstable because <1% of the cohort lived past 85 involving extrapolations of the fitted models beyond the range of the data upon which they are based (see Valdez-Flores et al. 2010 for detailed discussion).
- Used CDC WONDER (Wide-ranging Online Data for Epidemiologic Research) Database 1999-2022 instead of NCI SEER (Surveillance, Epidemiology and End Results) Database. The CDC WONDER is more comprehensive than SEER in that it covers 100% of the US through the National Program of Cancer Registries compared to 30% of the US population.

To address the US EPA's request for alternative values to the 2019 Memorandum, this analysis followed the US EPA IRIS (2016) approach for life-table calculations, modeling excess risk for lymphoid and breast cancer incidence based on the lymphoid and breast cancer mortality data from Steenland et al. (2004). However, as detailed by Sielken and Valdez-Flores (2009), this approach is flawed. To increase transparency regarding these limitations and uncertainty, EPA could limit calculations of excess when only mortality data is available, while applying an uncertainty factor (UF) based on the difference between mortality and incidence extra risk calculations.

Table 2. Summary of Biologically Plausible EtO Cancer incidence unit risk values based on Rat and Mouse Cancer and Human Cancer Risk (without ADAF) Values

Values are reported without ADAF to be comparable to the 2019 USEPA Memo

Cancer/Model Type Note: All estimates for lymphoid and breast cancer are based on NIOSH study by Steenland et al. 2004)	Model slope per ppm-day MLE (SE)	Incidence Unit risk per ppm MLE (95% UCL)[7]
Rat Cancer Risk (USEPA IRIS, 2016, Table 4-20, 4-26)	-	4.01 – 6.70 E-02 (95% UCL)
Mouse Cancer Risk (USEPA IRIS, 2016, Table 4-26)	-	5.51 – 8.33 E-02 (95% UCL)
Lymphoid (15 yr lag) + Breast Cancer (20 yr lag) from NIOSH Steenland et al. (2004) study converted to incidence unit risk (USEPA's method)	See below for each cancer type	
log-linear CPH (TCEQ, 2020 and Valdez-Flores 2010)		2.79E-02 (5.64E-02)
log-linear CPH (USEPA IRIS, 2016)		3.89E-02 (7.01E-02)
linear CPH (USEPA IRIS)		8.66E-02 (1.93E-01)
Lymphoid Cancer Mortality (15 yr-lag)		
log-linear CPH (TCEQ, 2020)	3.12E-06 (2.61E-06)[1]	1.93E-03 (4.59E-03)
log-linear CPH (USEPA IRIS, 2016)	4.74E-06 (3.35E-06)[2]	2.93E-03 (6.34E-03)
linear CPH (USEPA IRIS, 2016)	1.23E-05 (2.12E-05)[3]	7.59E-03 (2.91E-02)
Breast Cancer Mortality (20 yr-lag)		
log-linear CPH (Valdez-Flores et al. 2010; Supplementary tables with lagged-out slopes)	9.42E-06 (5.48E-06)[4]	2.78E-02 (5.43E-02)
-log-linear CPH (USEPA IRIS)	1.22-E05 (6.41E-06)[5]	3.60E-02 (6.70E-02)
-linear CPH (USEPA IRIS)	2.68E-05 (2.15E-05)[6]	7.90E-02 (1.83E-01)

[1]Table 9 in TCEQ (2020) – mortality – males only (higher cancer risk than for males and females combined)

[2]Table 4-2 in USEPA (2016)

[3]Table D-36 in USEPA Appendices (2016). Calculated SE=(4.71E-05-1.227E-05)/1.645

[4]Table S-10 in Supplemental of Valdez-Flores (2010)

[5]Table 4-9 in USEPA (2016)

[6]Table D-25 in USEPA Appendices (2016).

[7]POD = 1E-05 (Sielken and Valdez-Flores, 2009), risk at 70 yrs age, CDC Wonder Database 1999-2022 (https://wonder.cdc.gov/cancer-v2022.html) – data covering all 50 states and the District of Columbia and Puerto Rico provided by The Centers for Disease Control and Prevention National Program of Cancer Registries (NPCR) and The National Cancer Institute Surveillance, Epidemiology and End Results (SEER) Program, lymphoid based on combined male and female background incidence rates, breast cancer based on female background incidence rates.

1.5. Additional Model Slope Comparisons Based On NIOSH Breast Cancer Incidence, Updated NIOSH Breast Cancer Mortality, UCC + NIOSH Combined Cancer Mortalities

For comparison purposes only, the following model slope values for breast cancer incidence (USEPA IRIS) and breast cancer mortality (Kelly-Reif et al. 2026) were also estimated. Breast cancer incidence is not appropriate for quantitative risk assessment purposes because there are a substantial number of missing cases. Based on analysis of Steenland et al. (2003) if all women were interviewed, the analysis would have included 134 or more breast cancer cases in addition to 233, which became the data incorporated in the USEPA model[15]. The missing cases were due mostly to location problems of short term workers, i.e., those with lower cumulative exposures. This can result in serious concern about selection bias where proportionally more cases are found and interviewed among long-term workers, and fewer cases interviewed in shorter term workers.

For the updated breast cancer mortality study, the "unadjusted" linear CPH slope reported by Kelly-Reif et al. (2026) is nearly identical (and numerically smaller) than the USEPA IRIS (2016) linear CPH slope based on Steenland et al. (2004), indicating consistency in linear model results across datasets. These results do NOT indicate, as USEPA claimed, that the results of this update support USEPA's steep spline model. The "unadjusted" linear CPH model is identical to the USEPA IRIS (2016) application of the linear CPH model. Notably, neither USEPA IRIS nor Kelly-Reif et al. (2026) evaluated the log-linear CPH model, despite its application to the breast cancer mortality data in Steenland et al. (2004).

[15] Steenland et al. (2003) identified 319 cases. The SMR = 0.87. 233 (73%) of cases interviewed. SMR = O/E; 0.87 = 319/E, solving for E: E=367 so 319 were found (observed) but expected is 367. Therefore 367-319 =48 cases under-ascertained. Assuming no association with EtO, i.e., SMR = 1. Observed would have to be 367, same as expected to get SMR of 1.0. If EtO caused cases then more than 48 would be missing to get an SMR>1.0. Of 319 cases, only 233 (73%) were interviewed, so 319-233 =86 were lost at this phase of study. So at least 48 + 86 =134 more cases should have been included in the interview portion of the study.

Table 3. Additional Model Slopes for NIOSH 2003 Breast Cancer Incidence, NIOSH 2004 Breast Cancer Mortality, NIOSH 2025 Breast Cancer Mortality, and Combined UCC and NIOSH Cancer Mortalities, for Comparison Purposes Only

Cancer/Model Type (lag included unless otherwise noted)	Model slope MLE (SE) per ppm days
Breast Cancer Incidence	
log-linear CPH (USEPA IRIS, 2016)	5.4E-06 (3.5E-06)[1] all 9.5E-06 (4.1E-06)[1] interview
linear CPH (USEPA IRIS, 2016)	2.30E-05 (1.44E-05)[2] interview
Breast Cancer Mortality	
linear CPH (Kelly-Reif et al. 2025;Supplemental)	1.92E-05 (1.83E-05)[3]
Linear CPH (USEPA IRIS analysis of Steenland et al. 2004)	2.68E-05 (2.15E-05)[4]
UCC + NIOSH studies for Male and Female combined	
Log-linear CPH (no lag) (Valdez-Flores et al. 2010)	1.92 E-06 (2.08 E-06)

[1]Table 4-12 in USEPA 2016

[2]Table 4-13 in USEPA (2016). Calculated SE=(4.67E-05-2.30E-05) /1.645

[3]MLE and SE for the slope were calculated from the unrestricted linear model in Table S3 of Kelly-Reif et al. (2025). MLE = (1.07-1)/3650 = 1.92E-05, 95%UCL = (1.18-1)/3650=4.93E-05, SE=(95%UCL-MLE)/1.645.

[4] USEPA IRIS, 2016 Table D-25

1.6. Permitted Daily Exposures (PDE) Based On New And Existing Genotoxicity Data Compared With USEPA And TCEQ 1 /Million Risk Specific Concentrations

In the absence of experimental evidence to the contrary, mutagenicity induced by EtO as reviewed in detail by USEPA IRIS (2016) and Gollapudi et al. (2020) has been hypothesized as the default mode of action for its carcinogenicity. EtO is a relatively weak genotoxicant, both in the magnitude of the observed responses and the doses and exposure durations needed to elicit a significant response (Gollapudi et al. 2020). Recently, the recognition that genotoxicity could be used on its own merit for human risk assessment has led to approaches to derive a Permissible Daily Exposure (PDE) (ICPEMC, 1980) based on *in vivo* genotoxicity dose-response data to adequately protect humans from adverse health outcomes resulting from mutagenic effects. As described in greater detail below, the following table compares PDE's based on existing and recently completed inhalation genotoxicity studies with other 1 in million excess risk values.

Table 4 Comparison of PDE or 1 in million excess risk values

Reference	Endpoint	PDE or $1x10^{-6}$ risk	ppt EtO
Gollapudi et al., 2020	Genotoxicity	PDE	238
Gollapudi et al., 2026	Genotoxicity	PDE	4905
US USEPA IRIS (2016)	Human Cancer	$1x10^{-6}$ risk	0.1
TCEQ	Human Cancer	$1x10^{-6}$ risk	240
NTP (1987)	Mouse Cancer	$1x10^{-6}$ risk	20
Lynch et al. (1984)	Rat Cancer	$1x10^{-6}$ risk	13

The PDE value of 4,906 ppt, based on newly published *in vivo* mutagenicity dose-response data (Gollapudi et al., 2026), is approximately 20-fold higher than the TCEQ's one-in-a-million extra cancer risk level and nearly 49,000-fold higher than the USEPA's. This builds on Gollapudi et al. (2020), which established a PDE of 238 ppt based on previously published *in vivo* genotoxicity data. Comparisons with animal cancer studies further show that while TCEQ values are 12 to 18 times higher than these benchmarks, USEPA values are more than two orders of magnitude lower.

Collectively, these 2020 and 2026 PDE values provide converging evidence for the biological plausibility of the TCEQ's dose-response modeling. By prioritizing biological plausibility from the outset to select the log-linear CPH model, the TCEQ framework offers a more scientifically defensible basis for determining EtO cancer risk than the USEPA's approach.

1.7. Rigorous Testing Of IRIS Conclusions Based On EtO Exposures In Smokers

USEPA's gold standard science policy includes a mandate to be "skeptical" "of findings and assumptions" and "structured for falsifiability of hypotheses." These principles refer to conducting science with an aim towards independently testing findings. This should be particularly true of the risk estimates from the NIOSH study, given that USEPA predicts extraordinarily low levels of EtO, less than background or endogenously produced, of EtO have a cancer risk.

One way to assess the plausibility of the IRIS cancer risk estimate for lymphoid cancer is with studies of smoking and lymphoid cancer. As demonstrated in greater detail below, smokers have significant analytically quantified EtO exposures, and smoking-related exposures can be reliably estimated with pack-year smoking data available in smoking epidemiology data. The Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial, which also evaluated the risk of non-Hodgkin Lymphoma (NHL), is the largest U.S.-based cohort study that evaluated smoking and lymphoid cancer outcomes (Troy et al., 2010). Based on the pack-year smoking data, known concentrations of EtO present in cigarette smoke and varying levels of smoking intensity; EtO exposures were estimated for the subjects in the study, and the IRIS value applied

to estimate lymphoid cases. The estimate of lymphoid cases, assuming IRIS as a correct predictor of risk, were large enough that they would have been statistically detectable based on the large sample size of the study (>1 million person-years of follow-up). Yet, there was no statistically elevated level of lymphoid cancer in the study. In fact, non-smokers had slightly higher (non-statistically significant) incidence of lymphoid cancer. This key smoker reality check is yet another analysis that shows the implausibility of the IRIS value.

USEPA's response to lack of a lymphoid cancer signal for smokers (USEPA, 2024) implausibly posits, without presenting any evidence, that the absence of a lymphoid cancer signal in smokers is due to possible interactions between EtO and other carcinogenic agents in cigarette smoke that attenuate possible EtO lymphoid cancers. In addition, USEPA (2022) itself recommended that if EtO concentrations could be quantified in cigarette smoke, such data would be a valuable "forwards" validation opportunity of the EtO IRIS low-exposure cancer risk predictions. Although this point has been made before (ACC, 2023; Kirman et al., 2025), USEPAthe new analysis in this document provides greater quantitative rigor.

1.8. Selection Of Endpoints And Models For Risk Assessment Purposes Are Not Indications Of Causality

The proposed selection of cancer endpoints and statistical models in section 1.2 is grounded in the established regulatory risk assessment framework in which the objective is to be public health protective, while remaining scientifically reasonable. Under this paradigm, regulatory decisions may appropriately incorporate conservative assumptions in the presence of uncertainty, provided they remain biologically and scientifically defensible.

The human evidence alone is insufficient to conclude that EtO is a carcinogen because the available epidemiological evidence does not consistently show elevated cancer risk (IARC, 2008; Valdez-Flores et al. 2025; Marsh et al. 2019; Steenland et al. 2003). However, public health agencies, including USEPA, routinely act under conditions of uncertainty using evidence that is less stringent than formal causality criteria (e.g., Bradford Hill considerations), consistent with their preventive public health mandate.

This approach is consistent with the National Research Council's 1983 Red Book (*Risk Assessment in the Federal Government: Managing the Process*[16]) which explained the use of inference judgments in the risk assessment process:

> "[A] desire to err on the side of overprotection of public health by increasing the estimate of risk could lead an assessor to choose the most conservative assumptions throughout the process for components on which science does not indicate a preferred

[16] National Research Council. 1983. *Risk Assessment in the Federal Government: Managing the Process*.

choice. Such judgments made in risk assessment are designated risk assessment policy, that is, policy related to and subservient to the scientific content of the process . . . ."

"To make judgments amid such uncertainty, risk assessors must rely on a series of assumptions."

For the purposes of establishing regulatory goals for EtO exposure, this document assumes an association between EtO and cancer, despite the inconsistency in the available evidence, particularly the epidemiologic evidence. In establishing recommendations for regulatory risk assessment levels, conservative assumptions were made that may overestimate the true risk, resulting in risk levels above the true risk, which may be zero, particularly at current exposure levels.

At the same time, risk estimates should not be overly conservative such that they are implausible. USEPA's *Guidelines of Carcinogen Risk Assessment* state (U.S. USEPA, 2005):

"In constructing high end estimates of risk, the assessor should bear in mind that the high-end risk is a plausible estimate of the risk for those persons at the upper end of the risk distribution (U.S. USEPA, 1992a). The intent of this approach is to convey an estimate of risk in the upper range of the distribution, but to avoid estimates that are beyond the true distribution. Overly conservative assumptions, when combined, can lead to unrealistic estimates of risk."

Accordingly, the risk estimates developed in this document are intended to overestimate the likely true risk value (e.g., by assuming a linear low-dose extrapolation) while avoiding model forms or assumptions that would produce biologically implausible exposure-response relationships resulting in scientifically implausible risk assessment.

2. Question 3: Is there new information that could be used in dose-response modeling, such as epidemiological studies of cancer in humans exposed to EtO?

Recent analysis of worker exposure information from the NIOSH study indicates there is a high probability for substantial misclassification of worker exposures prior to 1978 (Bogen et al. 2019). A major concern is that the NIOSH exposure model predicted that exposures were lowest in the 1940s and peaked in 1978—a trend that contradicts historical industrial hygiene realities such as poorer ventilation, fewer or no air washes before opening the chamber doors, poorer door seals, leaky gaskets, and weaker controls in earlier decades. In other words, workers with lymphoid cases who worked in earlier years and who in fact had higher exposures are most likely to be misclassified as having lower exposures. New analysis of Valdez-Flores et al. (2010) supplementary tables indicate that the exposures for the majority of NIOSH lymphoid cancer cases across all cumulative exposures occurred prior to 1978. Thus the substantial exposure misclassification for majority of cases precludes use of the NIOSH study for quantitative risk assessment purposes. This applies not only to Steenland et al. (2003, 2004) used by both USEPA IRIS and TCEQ as critical studies for cancer risk assessment; but also recent updates of the NIOSH study (Kelly-Reif et al. 2025; Picciotto et al. 2026).

The Valdez-Flores et al. (2025) update of the Union Carbide Corporation (UCC) cohort provides a robust, independent weight-of-evidence challenge to the USEPA's 2016 IRIS two-piece linear spline model by leveraging comparable statistical power for the most critical endpoint. Despite smaller total cohort size, the UCC study's 40-year follow-up yielded nearly identical male lymphoid cancer deaths (25 vs. 27 in the NIOSH cohort), providing equivalent power to detect the high-potency elevation in risk claimed by the USEPA. Given that UCC workers had more than double the average cumulative exposure (67 ppm-years) of the NIOSH cohort, the persistent absence of an increased cancer signal across multiple updates serves as a powerful refutation of the USEPA's steep dose-response model. Furthermore, the UCC study is grounded in actual plant monitoring data starting in 1957, unlike the NIOSH model's reliance on backward extrapolation based on almost no monitoring data prior to 1978. While the USEPA has suggested that "other chemicals" in the UCC plant might confound these results, this argument is logically flawed. There are no known chemicals in the specific UCC units studied associated with an increased risk of lymphoid tumors. For other chemicals to "hide" a real EtO-related risk, they would need to have an anti-carcinogenic effect, which is not biologically plausible.

The USEPA's recent assertion to the National Academy of Sciences (NAS)—that the Kelly-Reif et al. (2025) NIOSH update validates the IRIS two-slope spline model—is technically premature and overlooks significant model selection bias. The 2025 update omitted the standard log-linear CPH model, which was the primary model in the Steenland et al 2004 study, thereby precluding a direct statistical comparison of fit. Furthermore, the Kelly-Reif et al. (2025) main analysis introduced a "duration of employment" adjustment not found in the earlier Steenland et al. (1991,

2003, 2004) publications. This adjustment creates a significant collinearity bias because duration is an intrinsic component of, and thus expected to be substantially correlated with, the cumulative exposure metric (the product of duration and intensity). In other words, adjusting for duration constitutes a statistical overadjustment that can artificially skew the observed risk pattern, a source of bias that is very well recognized (Mohner and Wendt, 2017).

Picciotto et al. (2026) utilized the recently updated NIOSH mortality data to investigate the presence of a Healthy Worker Survivor Effect (HWSE), arguing that bias associated with the HWSE caused previous studies to underestimate EtO risk. However, Kelly-Reif et al. (2025) observed no risk attenuation with longer follow-up—a hallmark of HWSE—attributing this to the long latency periods of solid cancers. These two papers contradict each other despite using the same breast cancer mortality data and having overlapping authors. While Picciotto et al. reported lower Hazard Ratios (HRs) among long-tenured workers as evidence of HWSE, these results are statistically non-significant and potentially compromised by exposure measurement errors.

Technical inconsistencies also raise serious concerns regarding data quality. The authors report a large, unexplained discrepancy between mean employment duration (10.7 years) and mean exposure duration (6.4 years) in a cohort where unexposed workers were specifically excluded. Additionally, their premise that “healthier” workers stayed longer despite EtO irritation is questionable because such irritation requires concentrations not seen in these plants for over 50 years. Ultimately, while the authors correctly note that standard duration adjustments (like those in Kelly-Reif et al. 2025) are inappropriate, they fail to employ the "g-methods" necessary to prove or correct the bias they allege.

2.1 New Information Regarding the NIOSH Study Exposure Assessment Quantitative Risk assessment.

2.1.1 New Information from Bogen et al. (2019)

The EtO dose-response model US EPA IRIS (2016) selected to extrapolate increased EtO-induced lymphoid cancer mortality risk was a model it fit to NIOSH occupational cohort study data on breast and lymphoid cancers experienced by 18,235 workers exposed to EtO for at least 3 months in 14 industrial EtO-sterilization plants located in 12 US states (Steenland et al., 1991, 2003, 2004). The NIOSH study was judged by US EPA to be a “high-quality” study based on availability of individual worker exposure estimates from a “high-quality” exposure assessment US EPA IRIS 2016 p. 3-68, note 13). This NIOSH exposure assessment used a comprehensive job-exposure matrix with exposure-level estimates from a regression model that incorporated a variety of plant and production variables and that explained 85% of the variability in an independent “validation” data subset, comprising 20% of all EtO measurements made.

Supporting that “high-quality” study characterization, US EPA IRIS (2016, Appendix A.2.8 at p. A-14) specifically cited the study’s exposure assessment and “verification” procedures described

by Greife et al. (1988) and Hornung et al. (1994).These NIOSH co-authors developed a regression model using 80% of the all time-weighted average measures of EtO exposure in workers' personal breathing zones acquired from the NIOSH EtO cohort workers between 1976 and 1985. The US EPA "high-quality" characterization of the NIOSH cohort study has been repeated subsequently, most recently in US EPA (2024, at p. 46), which also notes at p. 47 that "the NIOSH regression model for estimating EtO exposure included data not only on job/work zone, but also on variables such as size of sterilizer, type of product, freshness of product, and exhaust systems for sterilizers." However, the model variables included in the NIOSH historical EtO exposure model notably omit two variables well recognized by industrial hygiene engineers to be critical drivers of occupational inhalation exposure concentrations, namely workplace volume and the rate of dilution by exchange with outside air (Bogen et al., 2019), despite the fact that such information concerning EtO sterilization work environments was available to US EPA who funded the Goldgraben & Zank, 1981) data cited by Bogen et al. (2019). Bogen et al. (2019) describe additional fundamental exposure data limitations of the NIOSH cohort study as follows

> "After reviewing a limited set of [NIOSH cohort exposure regression] model predictions showing an ***increasing*** trend in [EtO] levels during 1938–1978, an USEPA Science Advisory Board [SAB] concluded that such behavior was "unlikely" [US EPA, 2015]. It was well recognized by others at NIOSH that before the late 1970s that [EtO] exposures among sterilizer workers "were likely to have been ***higher***. . . before installation of engineering controls, when the OSHA standard was 50 ppm" [Steenland et al., 1991]. Industrial hygiene data collected in other industries indicate historical occupational exposure levels as well as industrial exposure for nearly all chemical toxicants show generally decreasing historical trends in exposure concentrations since the establishment of corresponding industrial exposure guidelines, such as the American Council of Government and Industrial Hygienists (ACGIH) threshold limit value (TLV) guidelines and regulatory limits such as OSHA regulations [see Bogen et al., 2019, Figure 1]. For example, such trends are well documented for industries using perchloroethylene and/or trichloroethylene [9–11]. Such decreasing historical reflect improvements in technology sophistication, operational efficiency, knowledge concerning toxicological effects, and the implementation of workplace of EtO, such trends are reflected by the decrease in ACGIH 8 h TLVs for respiratory EtO exposure over guidance values and standards. In the case of EtO, such trends are reflected by the decrease in ACGIH the period 1948–1984 [12] (see Table 1)." [emphasis added]

The good post-1978 NIOSH EtO-exposure model fit was driven largely by its calendar-year variable, which was considered "a surrogate for improvement in work practices due to increased awareness of the potential health effects of ETO" (Hornung et al. 1994). However, the calendar

year assigned in the model was arbitrarily set to 1978 for all years prior to 1978. In other words, conditions in earlier years (1940's – 1960's) were assumed to be the same as for 1978. However, there is well documented evidence that this is an incorrect assumption. For example, sterilizer air/nitrogen wash cycles that reduced EtO released into workplace air gradually increased from earlier years. In the earlier periods of the cohort, workers spent more time near newly-treated and stored products that off-gas EtO. Also, there was less ventilation and leakier sterilizer-unit seals during earlier years, and no pre-1976/1978 EtO exposure estimates were actually validated because data allowing such validation were unavailable for the NIOSH cohort during that period.

Steenland et al. (2004) noted that within the NIOSH study cohort, "Exposure levels generally diminished sharply in the early 1980s after the reports of a haematopoietic cancer effect in animals and humans." The NIOSH regression-model assumptions are inconsistent with historic worker exposure guidance, including ACGIH standards of 100 ppm in 1948 and 50 ppm in 1947) and published literature indicating important work practice changes prior to 1978 (Bogen et al., 2019; Bruch, 1972; USFDA, 1978; Goldgraben and Zank, 1981; Stetson et al., 1976; White, 1977). In contrast to expected greater workplace EtO concentrations prior to 1978, the NIOSH-cohort exposure model predicted that EtO concentrations experienced by the 95th-percentile-exposed workers who died of lymphoid cancer *increased* substantially (rather than decreased) starting from 1970 (US EPA IRIS, 2016, Figure D-23, p. D-67 lower "Cases" panel), similar to the exposure trend implied for relatively highly exposed lymphoid-cancer cases (Bogen et al., 2019, Figure 1).

It is such a commonly accepted concept that exposures in industrial environments were lower in the early part of the 1900s compared to later periods, particularly when regulations started expanding in the 1970s, that the National Academies of Sciences (NAS) just assumed that NIOSH estimated higher exposures in the earlier period of the cohort:

> "The committee also notes that the highest occupational exposures in the NIOSH cohort occurred during earlier time periods when exposure assessment was less robust (prior to extensive personal monitoring) …" (NAS, 2025).

However, application of the NIOSH exposure model prior to 1978 erroneously predicted lower exposures in the earlier time period when most of the workers who died of lymphoid cancer were exposed.

2.1.2 New Analysis of NIOSH Exposure Data

A new analysis of EtO exposures in the NIOSH cohort among 53 lymphoid cancer cases is summarized in Figure 1. This Figure shows that vast majority of the exposure for lymphoid cases

occurred in the pre-measurement period where USEPA erroneously assumed lower, not higher, exposures than during later period addressed by IH measurements.

For each case, years worked are represented as a horizontal line segment.

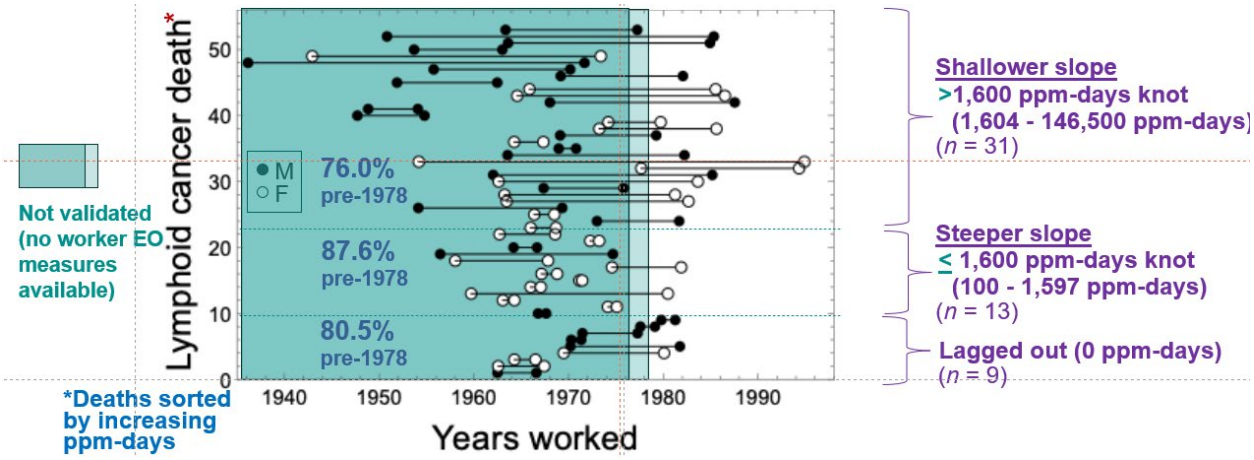
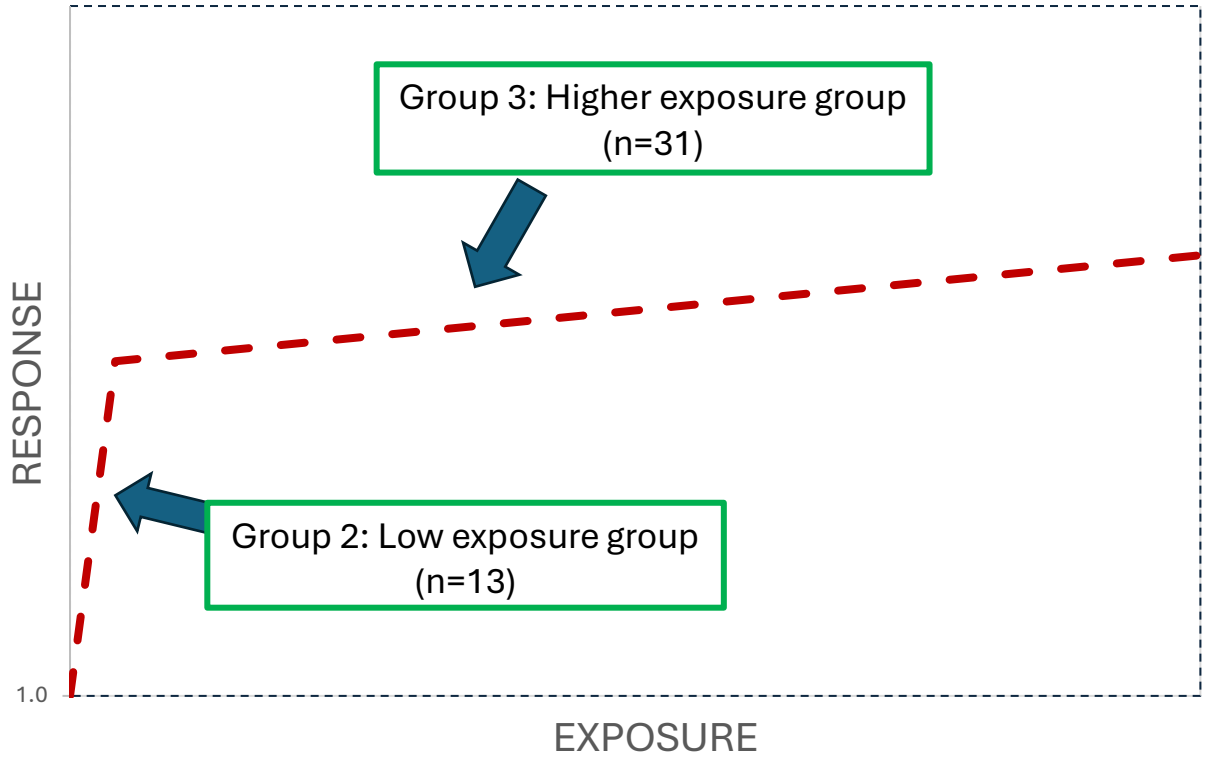


Figure 1. Range of employment years (x-axis) for 53 NIOSH cohort workers who died of lymphoid cancer (cases indexed on the y-axis), ordered by increasing cumulative exposure within three exposure categories: lagged-out (zero effective exposure), ≤1,600 ppm-days, and >1,600 ppm-days. Shaded regions indicate periods with no (dark shading) or limited (light shading) exposure measurement data. Male and female cases are denoted by closed and open symbols, respectively. Source: New analysis of summary data cited by Valdez-Flores et al. (2010).

Cases in Figure 1 are grouped into three exposure categories consistent with the USEPA IRIS (2016) two-slope spline model reflecting increasing cumulative exposure (see Figure 2):

- **Lagged-out cases (n = 9):** assigned zero effective exposure under a 20-year lag structure;
- **Low-exposure group (n = 13):** cumulative exposure ≤ 1,600 ppm-days (below the spline knot); and
- **Higher-exposure group (n = 31):** cumulative exposure > 1,600 ppm-days (above the spline knot).

Figure 2: Graphical Representation of the USEPA IRIS (2016) 2-Piece Linear Spline CPH Model Cancer Risk Assessment for Ethylene Oxide.



Note: Group 1 is not shown because it is the group of workers for which <u>all</u> of their exposure occurred within the lag period of 20-years. These workers are assigned zero effective cumulative exposure in the risk model even though they were exposed to EtO.

Across these three groups, the majority of cumulative exposure is based on modeled rather than measured data. Specifically, 80.5%, 87.6%, and 76.0% of cumulative exposure for the lagged-out, low-exposure, and higher-exposure groups, respectively, lack corresponding measurement data (Figure 1). In total, more than 75% of the cohort's cumulative exposure estimates are not supported by any direct exposure measurements.

The steep initial slope of the USEPA IRIS (2016) two-slope spline model—which largely determines the estimated potency for EtO-induced lymphoid cancer—is derived from the low-exposure group (n = 13). For this critical group, nearly 90% of cumulative exposure is based on entirely estimated rather than on any measured values. In addition, historical industrial hygiene evidence cited by Bogen et al. (2019) indicates that exposures in earlier time periods are likely underestimated in the NIOSH exposure model. These factors introduce substantial uncertainty

into the exposure–response relationship and directly affect the reliability of the resulting risk estimates.

This pattern is similar to that for breast cancer mortality cases (note: breast cancer incidence data are not available from NIOSH). As with lymphoid cancer mortality, the large majority of exposures to cases are based on little or no actual exposure data. Respectively, 91%, 90% and 91% of exposures to breast cancer mortality cases who are lagged out (n=42), are below the knot of 700 ppm-days used by USEPA IRIS (2016) (n=15), and above the knot (n=45) are based on exposures prior to 1978 (Valdez-Flores et al. 2010 supplementary summary tables).

Given the significant uncertainty in exposures, relying on biologically plausible models is crucial

2.2 New Epidemiology Studies

2.2.1 Valdez-Flores et al. 2025

Two EtO cancer risk assessments based on the NIOSH study, USEPAs two-piece linear spline model with a steep initial slope and TCEQs standard Cox proportional hazard (CPH) model, result in vastly different risk estimates. Valdez-Flores et al conducted analysis using an update of the Union Carbide cohort (UCC) to examine the epidemiological evidence for the shape of the dose-response model for EtO. The authors support use of this study of male chemical workers given the high average cumulative exposure (67 ppm years), extensive average follow-up of 40 years and the number of male lymphoid cancer deaths (25) comparable to the number observed in the NIOSH study.

The authors conclude that this independent analysis of a different cohort with comparable statistical power and cumulative exposures to the NIOSH study provides no empirical support for an initial steep slope. The UCC study of males was unable to address breast cancer but is informative with respect to the shape of the dose-response model based on the strongest evidence (male lymphoid tumors) reported by the Steenland et al. 2004 mortality study of the sterilant workers and which comprises 85% of the IRIS risk estimate.

2.2.2 Kelly-Reif et al. 2025 – Update of Steenland et al. (2004) for breast cancer mortality

Kelly-Reif et al. (2025) present an updated analysis of breast cancer mortality in the EtO cohort originally reported by Steenland et al. (2004). Unlike the earlier publication, the update does not re-evaluate lymphoid cancer mortality, which contributes approximately 85% of the USEPA IRIS cancer slope factor derived from the NIOSH cohort. Lymphoid cancers, which Steenland et al. 2004 and IRIS 2016 focus upon, are nowhere mentioned.

The Kelly-Reif paper presents results for breast cancer mortality only, using the log-log model with a 20-year lag[17], and did not address LH cancers at all. The related Picciotto et al. (2026)

[17] subjects (cases and non-cases) were assigned to an assumed zero exposure referent group if all their exposures occur within 20 yr. from end of their follow up.

study discussed in greater detail below, addressed LH and breast cancer mortality, but also not lymphoid cancer mortality.

Kelly-Reif et al. (2025) conclude that EtO exposure–response relationship is steep at low cumulative exposure levels. This conclusion is driven by several modeling and analytic choices, including:

1. Use of a log–log model, which applies a log transformation to cumulative exposure and can induce a more-than-linear (supra-linear) steep curvature;
2. Adjustment for duration of employment despite cumulative exposure already incorporating duration, raising concerns regarding well-recognized bias introduced by over-adjustment for duration (Mohner and Wendt, 2017)
3. Truncation of data in the high cumulative exposure range, which modifies the shape of the exposure-response curve, increasing steepness in the low exposure range.
4. The analysis does not present results using the standard CPH model, which served as the primary model in the original NIOSH analysis (Steenland et al. 2004).

Standardized mortality ratio (SMR) analyses using the same 20 yr lag do not suggest a steep slope in the low cumulative exposure range. Most exposure categories show non-statistically significant increases in SMR. The referent group shows a deficit (0.83), which may partially explain the reported risk increases in the internal comparisons. The authors dismiss the SMR results and assume these findings are due to a healthy worker survivor bias (HWSB).

HWSE may be present when long duration workers, who typically have higher cumulative exposure, are healthier with respect to the disease being studied than the comparison group. This can introduce downward bias in risk estimates (i.e., HWSB). In such cases, the HWSB may mask or attenuate the relationship between exposure and disease. However, this interpretation cannot be assumed for studies of female workers and breast cancer. In this context, long duration employment (greater cumulative exposure) may be associated with fewer or no pregnancies that would be associated with increased not decreased risk of breast cancer. A more likely explanation is the opposite from Kelly-Reif's assumption, namely an "unhealthy" worker survivor effect for breast cancer mortality. The authors did not consider that the deficit in the referent group (0.83) could be due to inclusion of greater numbers of shorter duration employees, who would typically have more pregnancies and lower breast cancer risk than the general population or than long term employees. This explanation may also be relevant to the consistently elevated RRs in the categorical analyses (Table 3 of Kelly-Reif et al.), together with the inappropriate control for duration of employment, which is complicated by duration being a component of cumulative exposure.

The study data related to parity (number of pregnancies) is supportive of a positive bias (the appearance of greater risks) and lack of attenuation at high exposures. In conjunction with the interview portion of the study, the authors examined parity as a potential confounder based on its known relationship with breast cancer (i.e., fewer pregnancies, greater risk for breast cancer) as

noted above. Kelly-Reif et al, also found that nulliparous women in the study had higher cumulative exposures, which confirmed parity as a potential confounder. However, the authors conclusion that "the HWSB likely causes bias towards the null" is not supported by their own data. Long duration working women (i.e., greater cumulative exposure) are likely to have fewer children and therefore a greater risk for breast cancer, inconsistent with bias towards the null. Although the authors state that controlling for potential confounders, including parity, did not result in "important changes in estimates that would alter the conclusion," their adjustment for parity is fairly rudimentary (yes or no births), and therefore, residual confounding may still be present.

In addition to the issues described above related to controlling for duration of employment, there is the added complication of measurement error of this variable. The authors failed to collect occupational data for this population from 1998 to 2021, the end of follow up. For workers still employed in 1998, the close of the Steenland et al. 2004 follow up, there is no information about when or if they left employment between 1998 and 2021. There is no explanation of how they treated duration of employment for this group of workers, except to say they were few in number and "inconsequential" with their 20-year lag. This absence of concern by the authors could have been tested, assuming various possible termination dates of those active in 1998. Errors in measurement of duration could impact both the estimation of cumulative exposure and the primary analyses that adjusted for duration.

The authors claim that their positive findings would be stronger had they relied on incidence data as survival from breast cancer substantially improved during the study period. In fact, the contrary may be true and a spurious positive association might manifest in the study if women exposed to higher levels in the earlier years also had lower survival.

Finally, the validity of this study and its conclusions are fundamentally related to the accuracy of the NIOSH study exposure estimates. We have described these issues in detail in our past comments, summarizing the implausible regression model exposure estimates for pre-1978 exposures that indicate lower exposures in the early years that increase rather than decline to 1978. Importantly, this potential measurement error impacts most of the study population who were hired prior to 1978, the time period when the regression model was not validated. Rather than expressing concern over the accuracy of exposure estimates prior to 1978, the authors truncate analyses at 5000 and 1000 ppm days, speculating misinformation at higher exposures.

Kelly-Reif et al. (2024) truncated the exposure data to apply their model results to occupational exposure limits, suggesting increased breast cancer mortality risk at the commonly adopted 1ppm and even at the more protective 0.1ppm. Using this process, their model showed increased risks at 365 ppm-days (equivalent to 10 years at 0.1ppm) and 3650 ppm-days (equivalent to 10 years at 1ppm) cumulative exposure. This argument, however, is misleading. As the authors point out themselves, a cumulative exposure of 365 or 3650 ppm days may be accumulated in many ways, since exposure is the product of duration and intensity. The product of shorter duration and

high exposure could result in these same ppm days. There is no evidence that low level exposure is actually driving these results.

2.2.3. Picciotto et al. 2026 Healthy Worker Survivor Effect in the updated NIOSH EtO study

Picciotto et al. (2026), co-authored by Kelly-Reif, utilize updated NIOSH mortality data for LH and breast cancer to argue that a Healthy Worker Survivor Effect (HWSE) has caused previous studies to underestimate true risk of EtO. Their analysis of the existence of a HWSE rests on two primary tests:

1) examining whether disease risk decreases with longer employment duration (suggesting healthier subjects work longer) (Table 3 of Picciotto et al. 2026) and

2) determining if higher exposure levels correlate with earlier employment termination (Table 4 of Picciotto et al. 2026).

Table 3 of the manuscript presents Hazard ratios (HR) overall and for quartiles of cumulative employment duration. The authors focus on whether there is a downward trend of disease with duration of employment (i.e., healthier subjects work longer). For both LH and breast cancer mortality there is no increased HRs overall or for any duration category, but the longest duration groups have the greatest deficits. However, none of the HRs are statistically significantly different from zero risk nor are they statistically different from the referent category. In addition, these analyses were adjusted for cumulative exposure up to the previous year, which is likely highly correlated with duration of employment, possibly resulting in overadjustment bias.

Table 4 of Picciotto et al. (2026) examines HRs for employment termination associated with exposures during the year prior to termination. Every cumulative exposure quartile for both diseases shows a statistically significantly increased HR for employment termination with no evidence of positive or negative trends. This contradicts the assertion in the abstract of this paper that, "Workers exposed to higher levels of ethylene oxide were also more likely to leave work."

The authors interpret their results as demonstrating that there is a Healthy Worker Survivor Bias (HWSB) but that duration of exposure should not be adjusted for in exposure-response models, for the same reasons we discuss in our comments above for Kelly-Reif et al. (2025)[18]. They add that in order to test for bias, g-methods are needed ***which they did not perform***. Instead they just declare there is a HWSB that causes underestimation of risks when using standard regression modeling.

[18] ". . .models that condition on employment status or duration of employment are also subject to bias" (cite Picciotto page number)

Several factors create serious uncertainty in the interpretation and conclusions of these authors:

- As discussed in comments related to Kelly-Reif et al. 2025 paper, there is measurement error associated with duration of employment and exposure estimates, the fundamental variables in assessment of HWSE.
- The acute toxicity levels of EtO that might create irritation or avoidance are in the hundreds of ppms, intensity levels not seen in over 50 years. This is not consistent with the authors' assumption that healthier workers would be better able to handle EtO irritation, stay employed and consequently have higher cumulative exposures (HWSE).
- In addition, Picciotto et al. reports substantially longer mean duration of employment (10.7) than mean duration of exposure (6.4) (Picciotto et al. Table 2). These two values should be similar, since all the employees in this study were exposed, as described by Steenland et al. 1991, the first publication of this cohort: "Since ethylene oxide in the air circulates freely, workers in any area sharing the same air with either the sterilizers or recently sterilized products were likely to have had some exposure. Workers judged never to have been exposed to ethylene oxide were excluded from the study cohort." This discrepancy between duration of employment and duration of exposure seems unreasonably large and worthy of verification, since these variables play a major role in the authors' analyses. Further evaluation, including obtaining the study data and modeling output files, is needed before this study can be used to make any conclusions.

3. Question 4: Is there any new information relevant to dose-response model selection such as consideration of statistical analyses (knot), visual model fit (response to comments on the HON) , or biological plausibility?

3.1. New Information Relevant to Biological Plausibility of Dose-Response Model Selection

3.1.1. Background

USEPA (2005) Cancer Risk Assessment Guidelines state:

- "*Another problem occurs when a multitude of alternatives are presented without sufficient context to make a reasoned judgement about the alternatives*
- *"This form of model uncertainty reflects primarily the availability of different computer models and not biological information about the agent being assessed or about carcinogenesis in general."*

Similarly, the USEPA SAB (2015) emphasized that "any model that is to be considered reasonable for risk assessment must have an exposure response form that is both biologically plausible and consistent with the observed data."

USEPA's responses to comments claim that biological plausibility was considered in the USEPA IRIS selection of statistical models applied to lymphoid mortality and breast cancer incidence data from the NIOSH epidemiology study. However, USEPA IRIS (2016) only considered biological plausibility within the narrow confines of comparing USEPA's preferred 2-piece linear spline model with only two highly implausible statistical models: (1) the 2-piece linear spline model with knot at 100 ppm-days and (2) the log cumulative model. USEPA correctly rejected these two models as biologically unrealistic, too steep (i.e., "approaching infinity as exposures decrease towards zero"), no exposed cases below the knot, and/or "not biologically probable for a complex, multistep process such as carcinogenicity in a variable human population" (USEPA IRIS, pp 4-10, 4-16). However, USEPA IRIS (2016) then uses the statistical significance of these two biologically implausible models to justify selection of the 2-piece linear spline model with knot at 1600 ppm-days. Although the USEPA IRIS assessment includes summaries of genotoxicity, toxicology, epidemiology and toxicokinetics, there is virtually no integration of these important lines of evidence into the final quantitative dose-response model selection process.

In response to public comments regarding the USEPA (IRIS) lack of consideration of the biological evidence in the exposure response assessment, USEPA (2022) performed a highly subjective visual inspection of genotoxicity and cancer bioassay data to support their claim that

biological evidence cannot be used to inform biological plausibility. The USEPA (2022) evaluation involved (a) plotting the data as point estimates without error bars, (b) drawing a straight line between the response levels for the lowest and highest dose levels, and (c) declaring the exposure response to be supralinear or sublinear depending on whether the responses for the mid-dose levels visually appeared to be above or below the line. This visual inspection did not involve any consideration of statistical significance or evaluation of which data set and dose regimen is most relevant and useful to inform the shape of an exposure-response curve based on epidemiology data.

In the following sections, new information is presented on various levels of biological complexity to contextualize the biological plausibility or lack thereof of the two models used for EtO cancer risk assessment, i.e., the USEPA's two-piece spline model and the TCEQ's log-linear Cox-proportional hazard (CPH) model.

3.1.2. New Animal Studies Reported By Liu et. al. (2026) And Gollapudi et. al. (2026)

These authors examined the shape of the dose-response for early key events in the genotoxic mode of action (MOA) for EtO induced tumorigenicity to inform the selection of a biologically plausible dose-response model for cancer risk assessment. EtO is an alkylating agent that reacts with nucleophilic sites on DNA. It mainly forms N7-(2-hydroxylethyl)guanine adducts (>80%), but the less common O^6-hydroxylethylguanine is usually linked to EtO's mutagenicity. EtO has been previously reported to be genotoxic both *in vitro* and in *vivo*, inducing mutations and cytogenetic damage. However, it requires high concentrations and prolonged exposure to produce genotoxic effects, indicating a relatively weak mutagenic potency (Gollapudi et al., 2020).

The US USEPA (2016) and TCEQ (2020) both assessed cancer risk from EtO exposure in the same NIOSH cohort of sterilizer workers, assuming a genotoxic mode of action and a linear no-threshold model. However, the US USEPA's calculated risk value (1 in a million extra risk = 0.1 ppt = 0.0001 ppb) is about 1000 times lower than TCEQ's (0.24 ppb). One of the reasons for this difference is the use of different statistical models: USEPA used a 2-piece linear spline for dose-response, while TCEQ applied a standard log-linear Cox proportional hazard (CPH) model. Both models fit the data similarly, but determining which is more biologically plausible remains as a critical gap.

The new investigations of Liu et al. (2016) and Gollapudi et al. (2026) examined the dose-response relationship for the early key events in EtO's genotoxic MOA to assess whether log-linear CPH model used by TECQ or two-piece linear spline model used by the US USEPA is more biologically plausible for assessing EtO cancer risks. Liu et al measured DNA adduct formation in multiple tissues (lung, liver, bone marrow, and mammary gland) whereas Gollapudi et al. assessed the induction of apical genotoxicity endpoints. In these studies, male and female mice were exposed via whole-body inhalation for 28 consecutive days to 7 concentrations of EtO (0.05-200 ppm) to establish the shape of the dose-response for these endpoints. The

concentrations ranged from low parts-per-billion environmental exposures up to substantially higher parts-per-million levels typical of occupational settings.

The molecular initiating event for EtO-induced mutagenicity is its reaction at DNA's nucleophilic centers. In tissue samples collected following 28 days of exposure, Liu et al. (2026) found that the *N7*-(2-hydroxylethyl)guanine (N7-HE-G) adducts increased linearly with exposure at all EtO concentrations and in all the tissues examined, whereas pro-mutagenic *O^6*-hydroxylethylguanine (O^6-HE-dG) adducts appeared only at exposures ≥50 ppm despite the use of highly sensitive detection methods. Dose-response slopes for both adducts became steeper only at higher concentrations, likely when detoxification and repair systems were saturated. Notably, N7-HE-G showed a steeper slope at higher exposures rather than plateauing, as predicted by the USEPA IRIS 2-linear spline model. This disproportionate increase in adduct load matched predictions from the previously published PB-PK model (Fennel and Brown, 2001), which attributes elevated blood EtO levels to EtO-induced glutathione depletion at high exposures.

Gollapudi et al. (2026) assessed genotoxicity by measuring the frequency of *Pig-a* mutations in reticulocytes (RET) and mature red blood cells (RBC) on Day 28 of the exposure. Cytogenetic damage was evaluated by the erythrocyte micronucleus (MN) test in blood samples collected on Days 5 and 28 of the exposures. EtO was found to be a relatively weak genotoxicant with treatment-related increases in *Pig-a* and MN frequencies being observed primarily at 200 ppm. The shape of the dose-response for genetic damage was conservatively interpreted as being no more than a linear response with a single slope.

The salient findings from Liu et al. (2026) and Gollapudi et al. (2026) studies are highlighted below.

a) <u>EtO is a Relatively Weak In Vivo Genotoxicant</u>

The findings from the Gollapudi et al. (2026) study informing relatively weak in vivo mutagenicity of EtO are consistent with previous research in bone marrow (Recio et al., 2004) and lung tissues (Manjanatha et al., 2017) of Big Blue® mice. Mutations increased only at higher EtO exposures (≥100 ppm in bone marrow; 200 ppm in lung) and after prolonged exposure (48 weeks for bone marrow; 8 weeks for lung). LeBaron et al. (2012) reported no MN induction in B6C3F1 mouse erythrocytes exposed to 10–200 ppm EtO for 28 days. The relatively weak genotoxicity of EtO was also highlighted by Gollapudi et al. (2020) review of 40 *in vivo* studies.

b) <u>Relevance of the 28-Day Exposure Studies in Mice to Assess Risk following Longer Term Exposures</u>

The studies by Gollapudi et al. (2026) and Liu et al. (2026) using a 28-day exposure regimen are relevant for assessing longer-term risks associated with EtO exposure in animal and human contexts. A 28-day period is sufficient to reach steady state for the

induction of DNA adducts in mice, rats, and humans (Filser and Klein, 2018; Walker et al., 2000), as adduct levels reflect formation, repair, loss, and cell turnover. Therefore, exposures beyond 28 days are unlikely to yield significantly different results. Notably, in the Gollapudi et al. (2026) study, gene mutations and chromosomal damage were mainly observed at the 200-ppm exposure, yet a linear, no-threshold dose-response was conservatively assumed. It is highly unlikely that longer exposures would produce effects matching the pattern predicted by the US USEPA's 2-piece linear spline model. For genotoxic chemicals, mutations occur early in carcinogenesis rather than after extended exposure periods (Moore et al., 2008).

c) Relevance of Mouse Data for Human Cancer Risk Assessment

An essential consideration in using data from the mouse studies reported by Liu et al. (2026) and Gollapudi et al. (2026) is the value of data from mouse models to inform the selection of an appropriate model for human cancer risk assessment, particularly since the EtO cancer risk assessment relies on human epidemiological data. Evidence suggests that the DNA repair capacity in humans is comparable to, if not greater than, that observed in mice (MacRae et al., 2015). This similarity extends to critical DNA repair mechanisms, such as O^6-methylguanine-DNA methyltransferase (MGMT), a highly conserved enzyme involved in the direct repair of the pro-mutagenic O^6-HE-dG and other O^6-dG alkyl adducts (Roy et al., 1995; Klapacz et al., 2016).

Furthermore, physiologically based pharmacokinetic (PB-PK) modeling of EtO reveals that systemic blood concentrations following exposure to EtO at levels ≤100 ppm are similar in both mice and humans, indicating that EtO detoxification toxicokinetics are largely consistent across these species (Fennell and Brown, 2001). This finding supports the use of mice as a conservative and valid surrogate model for predicting the genotoxic effects of EtO in humans, especially at exposure levels that do not overwhelm detoxification pathways essential for maintaining homeostasis.

At higher exposure levels (>200 ppm), PB-PK modeling indicates that humans possess a greater capacity to maintain EtO detoxication, for example, through glutathione conjugation, than mice. Consequently, the mouse response at these levels serves as a worst-case scenario, further reinforcing the conservative nature of using mouse data for human risk assessment (Fennell and Brown, 2001).

d) No Evidence for a Linear Two-Piece Spline Dose Response

The dose-response for both DNA adducts and the apical genotoxicity endpoints is conservatively more consistent with a log-linear (nearly linear) or single-linear (Cox proportional hazards, CPH) dose-response than the two-piece spline dose response with a steeper initial slope.

3.1.3. Inference From Animal Studies (Kirman et. al., 2021)

Cancer bioassays conducted in rats show that high EtO concentrations (10 ppm - 100 ppm) significantly increase cancer incidences, while similar tests for ethylene at high concentrations (300–3,000 ppm) do not produce increased tumor rates. HEV measurements indicate that elevated ethylene exposure leads to much higher systemic levels of EtO (up to 160-fold higher) than the endogenous levels, yet no rise in tumors was observed. The robust study design of ethylene cancer bioassay and lack of increased cancer incidence suggest the biological implausibility of a steep initial linear slope as predicted by the USEPA's 2-piece spline model. The ethylene data are conservatively consistent with a shallower slope predicted by TCEQ's log-linear CPH model.

3.1.4. Inference From Human Data On Smoking And Lymphoid Cancers

Smokers are exposed to higher levels of EtO compared to nonsmokers, but their exposures remain lower than those seen in the most highly exposed workers. Data from smoking populations can help shape our understanding of EtO's dose–response curve and its associated cancer risks at low exposure levels.

Neither the Surgeon General's Report nor the American Cancer Association have linked lymphoid cancers to smoking, and evidence for a connection with breast cancer is weak. This suggests that low-level EtO exposures pose limited cancer risk, making a steep exposure–response model at these levels inappropriate.

The plausibility of the IRIS value can be directly tested using data from The Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial, which evaluated the risk of non-Hodgkin Lymphoma (NHL) and its subtypes in association with anthropometric factors, smoking, and alcohol consumption (Troy et al., 2010). PLCO is a large prospective cohort study by the U.S. National Cancer Institute. The key scientific question is whether the studies have enough statistical power to detect the lymphoid cancer incidence one would predict with the IRIS value.

PLCO includes estimates of NHL based on pack-years of smoking. The results are summarized in Table 5. The exposed subjects (smokers) were divided into four groups based on pack-years of smoking: (1) <8.75 pack years, (2) 8.75-22.4 pack years, (3) 22.5-39.75 pack years, and (4) >39.75 pack years. There were 564,061 person-years in the control group and between 148,440-161,046 person-years in the exposed groups. The hazard ratios comparing the incidence of NHL in the exposed groups to the control group ranged from 0.91-0.99 and were not statistically significant and <1 (fewer cases in exposure groups), consistent with no association.

Table 5. Hazard ratios in Troy et al. by pack-years of smoking

Pack-Years of Smoking	Person-Years	NHL Cases During 8.5 Year Follow-Up	Observed Hazard Ratio
Never	565,061	603	1.0
<8.75	161,046	152	0.91 (0.76-1.09)
8.75-22.4	156,415	173	0.99 (0.83-1.17)
22.5-39.75	151,920	158	0.94 (0.79-1.12)
>39.75	148,440	165	0.95 (0.80-1.14)

Applying the methods in Kirman et al. (2017, 2021) for estimating EtO exposure from biomonitoring measurements to estimate EtO exposures for each exposure group yields the following estimates:

- <8.75 pack years: 1.0 ppb
- 8.75-22.4 pack years: 3.6 ppb
- 22.5-39.75 pack years: 7.1 ppb
- >39.75 pack years: 11.5 ppb

Applying the IRIS risk estimate to the Troy et al. (2010) study is not straightforward. In particular, we need to account for the 8.8 year follow-up in the study, as opposed to estimating a lifetime risk. The following steps were taken to estimate risks for each exposure group, consistent with USEPA assumptions in developing the IRIS value:

1. The unit risk estimate (URE) is an extra risk per ppm (e.g., 3.69 per ppm for lymphoid in adult-only exposures Table 4-21 IRIS 2016) for an occupational career. This results in a lifetime exposure risk of 4.78 per ppm (i.e., 3.69*(70/54), see page 4-88 in IRIS 2016) for the upper-bound of the IRIS risk (the value applied for risk assessment) or 1.21 per ppm (i.e., 0.93*(70/54), see page 4-21 in IRIS 2016) for the maximum likelihood estimate (MLE)).

2. The extra risk per ppm of lymphoid for adult-only exposures (age 16 and older) for any period of time (ageEND-ageSTART) would be equal to 4.78*(ageEND-ageSTART)/70 per ppm (page 4-89 in IRIS 2016) for the upper-bound or 1.21*(ageEND-ageSTART)/70 per ppm for the MLE. Assuming risk accrues linearly with duration (i.e. a default assumption in risk assessment), this is equivalent to an extra risk of 4.78/70 ≈ 0.068 extra cases per 1,000 persons per ppm per year of exposure (or 0.068 per person-year per ppm) for the upper-bound and 1.21/70 ≈ 0.017 extra cases per 1,000 persons per ppm per year of exposure (or 0.017 per person-year per ppm) for the MLE.

3. The results of the analysis are shown in Table 6 (IRIS upper-bound) and Table 7 (IRIS MLE). Using the IRIS unit risk (upper-bound) for EtO, a 24-hour average EtO concentration of 11.5 ppb (corresponding to the high smoking group) yields an estimated lifetime excess lymphoid cancer risk of 7.9×10^{-4} per person. Applied to the 148,440 person-years in the high-exposure category, this implies approximately 117 additional lymphoid cases attributable to EtO using the upper-bound IRIS value, of which roughly 70 would be non-Hodgkin lymphoma (NHL) assuming NHL comprises 60% of lymphoid malignancies, as in the IRIS analysis.

    a. In the cohort data, 165 NHL cases were observed in this high-exposure category, with no elevation in risk compared to the reference group; these can therefore be reasonably considered background cases. If the IRIS unit risk (upper-bound) were correct, the expected total NHL count in this high-exposure group would be approximately 235 cases (165 background + 70 EtO-attributable). Given the observed person-time and events, this would correspond to a hazard ratio of about 1.5 (95% CI: 1.3–1.7), which would be clearly detectable and statistically significant. Because such an elevation is not observed, the IRIS unit risk is inconsistent with epidemiologic evidence at these exposure levels.

    b. The same reasoning applies to the next lower exposure group (22.5–39.75 pack-years), in which an IRIS upper-bound prediction would also imply a statistically significant elevation (estimated HR ≈ 1.25, 95% CI: 1.1–1.5), yet the study data do not show such an increase.

    c. Applying the MLE yields only 18 additional cases in the high exposure group compared to 165 observed. The estimated hazard ratio is 1.2 (95% CI: 0.98-1.4), which is not statistically significant. Thus, the cohort study is consistent with the MLE, though it does not prove or disprove that the MLE is accurate.

In summary, the smoker lymphoid cancer analysis shows that the Troy et al. cohort study is inconsistent with the IRIS upper-bound value (the value used for risk assessment). If the IRIS upper-bound value was the scientifically accurate lymphoid cancer risk for EtO, it would yield excess cancers in the Troy et al. cohort that could be detected given the statistical power of study, but no excess cancers were observed. The IRIS MLE estimate yields insufficient excess cancers to be detected in the Troy et al. study; thus, the PLCO study has insufficient power to test the hypothesis that the MLE IRIS lymphoid estimate is consistent with observed cases.

This analysis shows that the IRIS upper-bound estimate for lymphoid cancer is incorrect and should not be used.

Table 6. Predicted total cases and hazard ratios assuming the IRIS upper-bound value

Group (pack-years)	Exposure	Risk (per ppb)	Person-Years	Extra NHL Cases Due to EtO	Observed Cases	Predicted Total Cases	Hazard Ratio
Control	0	n/a	565,061	n/a	603	603	n/a
<8.75	1.0	6.9×10^{-5}	161,046	7	152	159	0.92 (0.78-1.1)
8.75-22.4	3.6	2.4×10^{-4}	156,415	23	173	196	1.17 (1.0-1.4)
22.5-39.75	7.1	4.8×10^{-4}	151,920	44	158	202	1.25 (1.1-1.5)
>39.75	11.5	7.9×10^{-4}	148,440	70	165	235	1.48 (1.3-1.7)

Table 7. Predicted total cases and hazard ratios assuming the IRIS MLE

Group (pack-years)	Exposure	Risk (per ppb)	Person-Years	Extra NHL Cases Due to EtO	Observed Cases	Predicted Total Cases	Hazard Ratio
Control	0	n/a	565,061	n/a	603	603	n/a
<8.75	1.0	6.9×10^{-5}	161,046	2	152	154	0.88 (0.75-1.1)
8.75-22.4	3.6	2.4×10^{-4}	156,415	6	173	179	1.07 (0.91-1.3)
22.5-39.75	7.1	4.8×10^{-4}	151,920	11	158	169	1.04 (0.88-1.2)
>39.75	11.5	7.9×10^{-4}	148,440	18	165	183	1.15 (0.98-1.4)

Beyond just lymphoid cancers, recent studies using NHANES/CDC data found no association between EtO exposure and cancer prevalence, including breast cancer (Kirman et al. 2025).

These findings indicate EtO exposure does not correlate with increased self-reported cancer cases for both smokers and nonsmokers.

Two key observations underpin low-dose risk assessments: (1) there is no clear link between cancer and EtO exposure in smokers; and (2) positive associations appear only among workers with much higher EtO exposures. Combined, human evidence aligns with dose–response patterns observed in animal studies where a steeper slope (cancer unit risk) is not observed at lower exposure levels.

3.1.5. Conclusions

Taken together, the above lines of evidence on the dose-response from four levels of biological organization (i.e., DNA adducts, apical genotoxicity endpoints, inference from animal carcinogenicity studies, and human epidemiology data from smokers) support the conclusion that the cancer risk to the general population from EtO exposure is conservatively modeled by assuming a MOA that is characterized by a shallow single-slope linear dose-response relationship. Overall, these data provided biological evidence in support of the TCEQ's log-linear model approach as a preferred and conservative approach for risk assessment of EtO carcinogenicity and shows the biological implausibility of the 2-piece linear spline dose-response model.

3.2. Derivation of Permitted Daily Exposures Based on the Genotoxicity Data from the 28-Day Inhalation Study in B6C3F1 Mice Exposed to Ethylene Oxide for 28-Days via Whole Body Inhalation (6 h/day, 7 days/week; Gollapudi et al., 2026)

3.2.1. ***Introduction:*** The following analyses were performed to establish a permitted daily exposure (PDE) value based on genotoxicity endpoints described in the study by Gollapudi et al. (2026) and to compare this PDE value with other relevant risk assessments. The use of genotoxicity dose response data to derive PDE values has been gaining significant attention in recent years (Heflich et al., 2020; White et al., 2020; Parsons et al., 2025). The Gollapudi et al. (2026) publication evaluated the induction of gene mutations (*Pig-a* mutations) and cytogenetic damage (micronuclei [MN]) in peripheral blood erythrocytes of male and female mice exposed to EtO via whole-body inhalation for 28 days.

3.2.2. ***Methodology:*** The methodology used by Gollapudi et al. (2020) was followed to identify the lowest PDE value based on the benchmark dose (BMD) modeling of the genotoxicity data from the Gollapudi et al. (2026) publication to identify BMDL50 values. The EU EFSA web-based program, which uses the R-package PROAST, version 70.0, was used for the BMD calculations (see Attachment). The following calculations were used to derive the values shown in Table 8 below.

Conversion to Continuous Inhalation (CI) Concentration: The estimated BMDL50 values were converted to continuous inhalation (CI) concentration (CIBMDL50) over 24 hr/day and 7 days/week by the following formula: CIBMDL50 = (BMDL50 x 6 x 7) ÷ (24 x 7).

Total Uncertainty Factor: Toxicodynamic (3.16) x Interindividual Variability (10) x Study Duration (10) x Severity of Effect (10) = 3160.

Permitted Daily Exposure: CIBMDL50 ÷ Total UF

3.2.3. ***Results:*** Only the endpoints reported in Table 8 were found to be amenable to BMD modeling. Please note that all models gave the following warning: *Attention: the AIC of the best model (minimum AIC) is more than two units larger than that of the full model. This might indicate a problem in the data, in particular when the difference is much larger than two units (e.g., > 5).* The warning may be the result of there being a response only at the top dose.

The lowest PDE value calculated from this analysis was 4905 ppt was derived from the dose response observed for the induction of *Pig-a* mutations in the reticulocytes of mice exposed to EtO. This value was used for comparisons with other risk values as shown in Table 2 below.

Table 8: Derivation of PDE values based on the genotoxicity endpoints reported by Gollapudi et al. (2026).

Endpoint	Parameter	BMDL50 (ppm)	CIBMDL50 (ppm)	UF	PDE (ppt)
Erythrocyte *Pig-a* Mutant Frequencies	Mutant RBC ($x10^{-6}$) – Females	71	17.8	3160	5,633
	Mutant RET ($x10^{-6}$) - females	62	15.5	3160	4,905
Erythrocyte Micronuculeus Frequencies	%MN-RET – females (28-day sampling)	149	37.3	3160	11,804

3.2.4. Comparison with previously derived risk values (Table 9):

The 1 in a million excess cancer risk value derived by TCEQ (2020) based on human epidemiology data was 240 ppt. The US EPA IRIS (2016) derived this risk value, based on the data from the same cohort, to be 0.1 ppt.

The inhalation unit risk (IUR) value for the mouse was 2.71×10^{-2} excess cancers per ug/m^3 or 4.88×10^{-2} excess cancers per ppm of EtO (NTP, 1987). This translates to a one in a million-cancer risk value of 20 ppt as calculated below:

$(1 \times 10^{-6}) \div (4.88 \times 10^{-2}) = 0.00002049$ ppm = approximately 20 ppt.

Similarly, the IUR for the rat is 4.17×10^{-5} excess cancers per ug/m^3 or 7.5×10^{-2} excess cancers per ppm of EtO (Lynch et al., 1984). This translates to a one in a million-risk value of 0.00001333 ppm or 13 ppt.

Gollapudi et al. (2020) have previously analyzed 40 sets of *in vivo* data sets reported in the literature on the genotoxicity of EtO and identified 238 ppt as the lowest PDE based on the induction of chromosomal aberrations in peripheral blood lymphocytes of mice exposed by inhalation to EtO.

Table 9. Comparison of PDE or 1 in million excess risk values

Reference	Endpoint	PDE or $1x10^{-6}$ risk	ppt EtO
Gollapudi et al., 2020	Genotoxicity	PDE	238
Gollapudi et al., 2026	Genotoxicity	PDE	4905
US USEPA IRIS (2016)	Human Cancer	$1x10^{-6}$ risk	0.1
TCEQ	Human Cancer	$1x10^{-6}$ risk	240
NTP (1987)	Mouse Cancer	$1x10^{-6}$ risk	20
Lynch et al. (1984)	Rat Cancer	$1x10^{-6}$ risk	13

3.2.5. Conclusions:

The PDE value identified based on the genotoxicity data reported by Gollapudi et al. (2026; i.e., 4905 ppt) was approximately 20-fold higher than the 1 in a million extra cancer risk value derived by the TCEQ and nearly 49,000-fold higher than the value for the extra cancer risk reported by the US EPA (Table 9). Gollapudi et al. (2020) have previously identified a PDE value of 238 ppt based on their analyses of published *in vivo* genotoxicity data. Similar comparisons of the risk values derived from animal cancer studies indicated that while the TCEQ value was 12- to 18-fold higher, the US EPA value was more than two orders of magnitude lower**. Collectively, the PDE values derived from the genotoxicity data (Gollapudi et al. 2020 and 2026) provide strong support for the biological plausibility of the dose-response modeling used by the TCEQ in deriving the EtO cancer risk.**

3.3. New Information Relevant To Statistical And Visual Fit And Consideration Of The "Knot"

Detailed comments by the ACC (2025) on USEPA's Proposed NESHAP Chemical Manufacturing Area Sources Technology Review provide recent comprehensive explanation for why USEPA IRIS calculation of p-values for USEPA's preferred 2-slope spline model are incorrect (see Table 1 above). This error led USEPA to select the 2-slope spline model and ignore the biologically more plausible standard linear or linear CPH models.

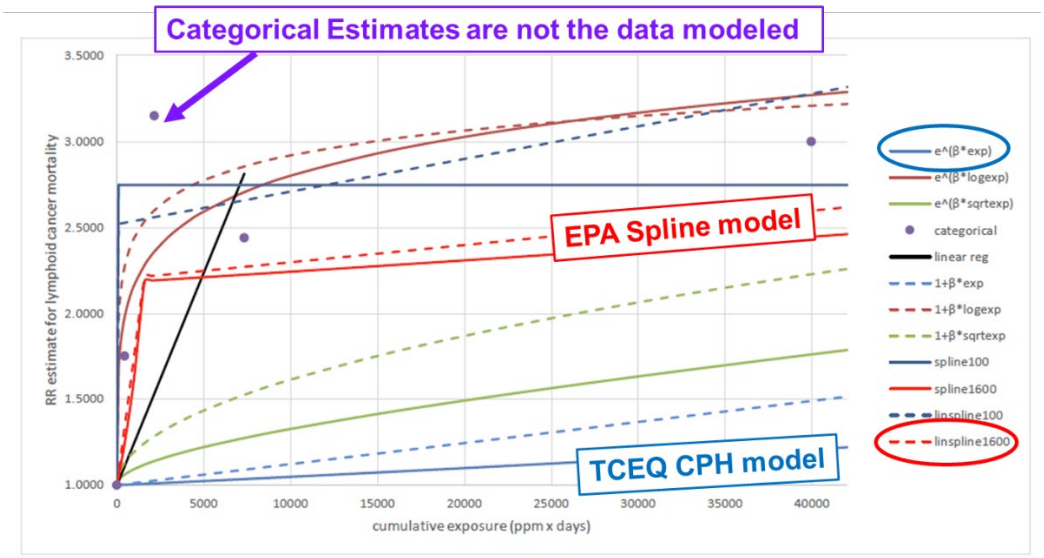
3.3.1 Background information on USEPA IRIS selection process for CPH models

USEPA IRIS (2016) evaluated a large number of different continuous models, each with a different equation. The continuous models are modeling the individual hazard rates and not the categorical estimates. Recently, USEPA has incorrectly equated continuous models and categorical estimates to all be based on the same comparisons leading to USEPA's incorrect comparisons of the models along the y-axis.

In USEPA IRIS Figure 4-3 shown below as Figure 1, the USEPA selected 2-piece linear slope spline model is the dashed red line and the solid blue line is the standard log-linear model. A spline model simply means two different models connected by a point of inflection, or knot. A 2-piece linear spline model means 2 different linear CPH models are connected by a knot. A 2-piece log linear spline model means 2 standard log-linear CPH models are connected by a knot. In both cases these models are based on individual hazard rates. The categorical results shown below as 5 purple dots are not the 53 individual hazard rates modeled.

USEPA IRIS (2016) evaluated many different types of CPH models to model these individual hazard rates. In an earlier draft USEPA IRIS, USEPA derived the cancer slope based on the categorical estimates. US EPA SAB reviewing this earlier draft rejected this method because the categorical (grouped) estimates often do not represent the pattern of the individual hazard rates modeled (Valdez-Flores and Sielken, 2013). However, recent US EPA RTC on the MON and HON reveal more clearly that US EPA selected their model based on incorrect comparisons of categorical estimates with continuous models both in general shape and numerically along the y-axis.

Figure 1. US EPA IRIS (2016) evaluated a large number of different CPH models, each with a different equation.



As shown in Figure 1, each type of CPH model has a different shape depending on the equation of the model used to fit the hazard rates.

A standard CPH model is a log-linear model, in which the relationship between the rate ratio (RR(d)) and the dose metric (usually cumulative exposure, e.g., ppm-days), "d," is log-linear as shown in Figure 2.

Figure 2. Standard log-linear CPH Continuous Model

Standard CPH Continuous Model

Relationship between the rate ratio (*RR(d)*) and the dose metric (usually cumulative exposure, e.g., ppm-days), *d*, is log-linear

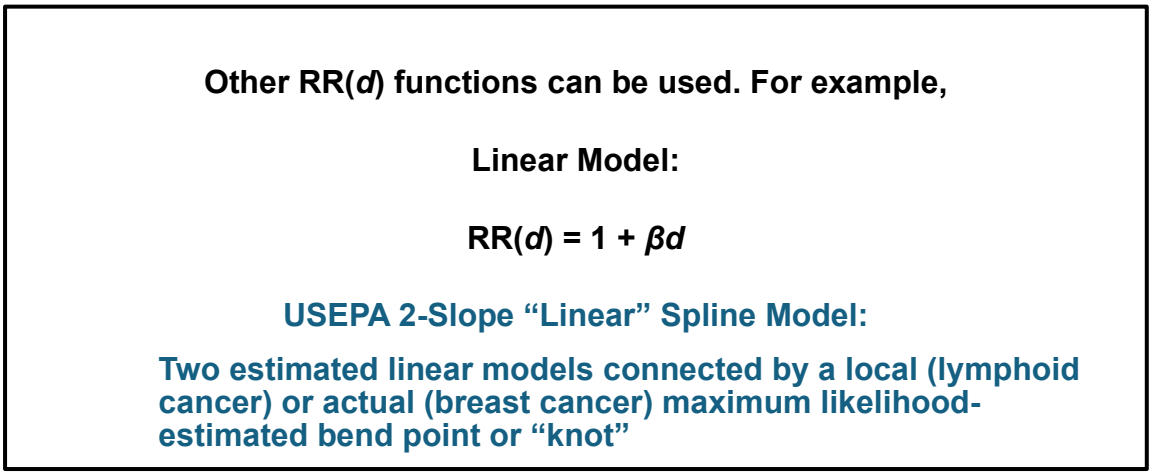
$$RR(d) = e^{\beta d}$$

β is estimated by fitting the individual hazard rates (HR)

Steenland et al. (2003, 2004) publications upon which IRIS is based focused mainly on standard log-linear CPH models using different exposure metrics (e.g., log cumulative or cumulative exposure). TCEQ (2020) also applied similar standard log-linear CPH models with cumulative exposure.

US EPA IRIS (2016) applied many other RR (d) functions. For example, the equation for the linear model is shown in Figure 3. US EPA's 2-slope "linear" spline model for lymphoid cancer mortality is based on two estimated linear CPH models connected by a local maximum likelihood-estimated bend point or "knot". Similarly, US EPA's 2-slope "linear" spline model for breast cancer incidence is based on two estimated linear CPH models connected by a overall maximum likelihood-estimated knot.

Figure 3. Linear and 2-slope linear CPH Model



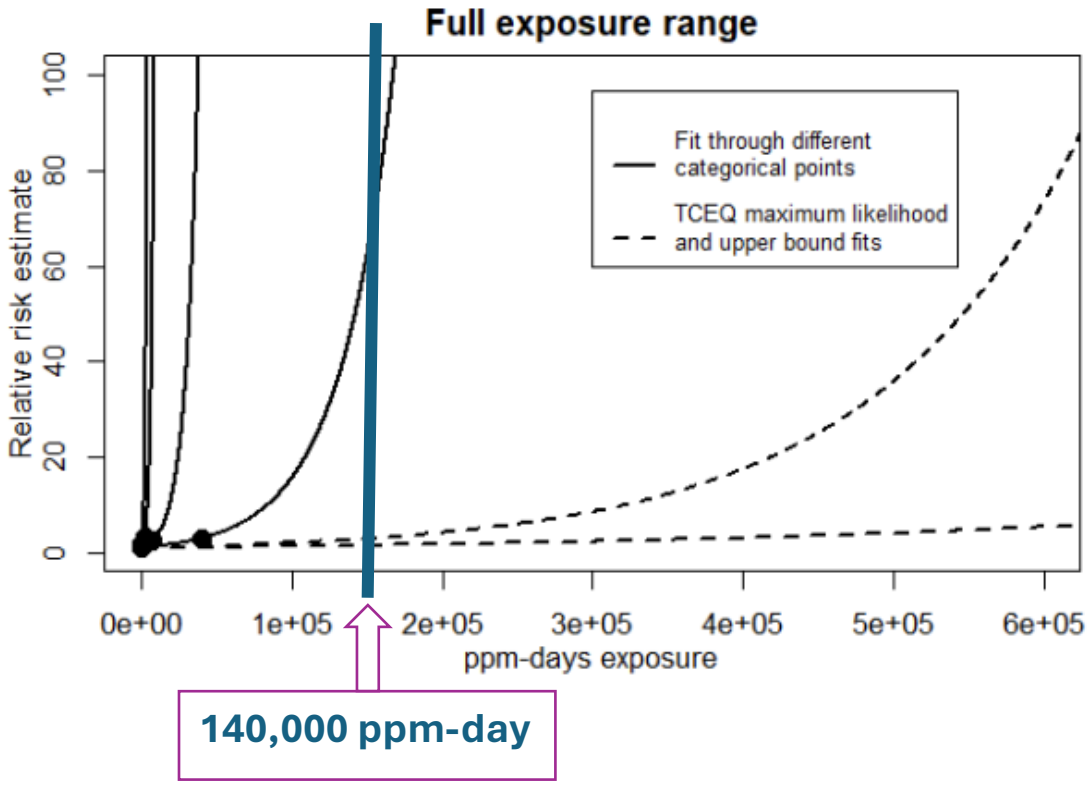
US EPA's RTC on the HON (p.49) incorrectly states:

". . . the log-linear Cox model is incapable of providing estimates of risk that are consistent with the categorical findings in the lower dose ranges of the NIOSH study **without "exploding" upward at the high doses received by some other workers in the study.**"

"That is this model is incapable of representing the increased cancer risk estimated with the categorical analyses at lower dose ranges **without making entirely unreasonable risk predictions at high doses."**

US EPA repeated these incorrect statements to the NAS reviewing the TCEQ (2020). Similarly, OEHHA repeated these claims to the California SRP reviewing OEHHA indicating that the standard CPH model "starts low and swings up high" and that the Standard log-linear CPH model "underpredicts the NIOSH data." This incorrect claim originated with US EPA (2022, p. 51) RTC on the MON which is an extreme misrepresentation of the TCEQ's model.

Figure 4. US EPA analysis used to reject the standard log-linear CPH model by erroneously fitting each categorical estimate separately and misrepresenting the full exposure range as 600000 instead of 140,000 ppm-day. The highest exposure range of the workers who died of lymphoid cancer is 146,460 ppm-day.



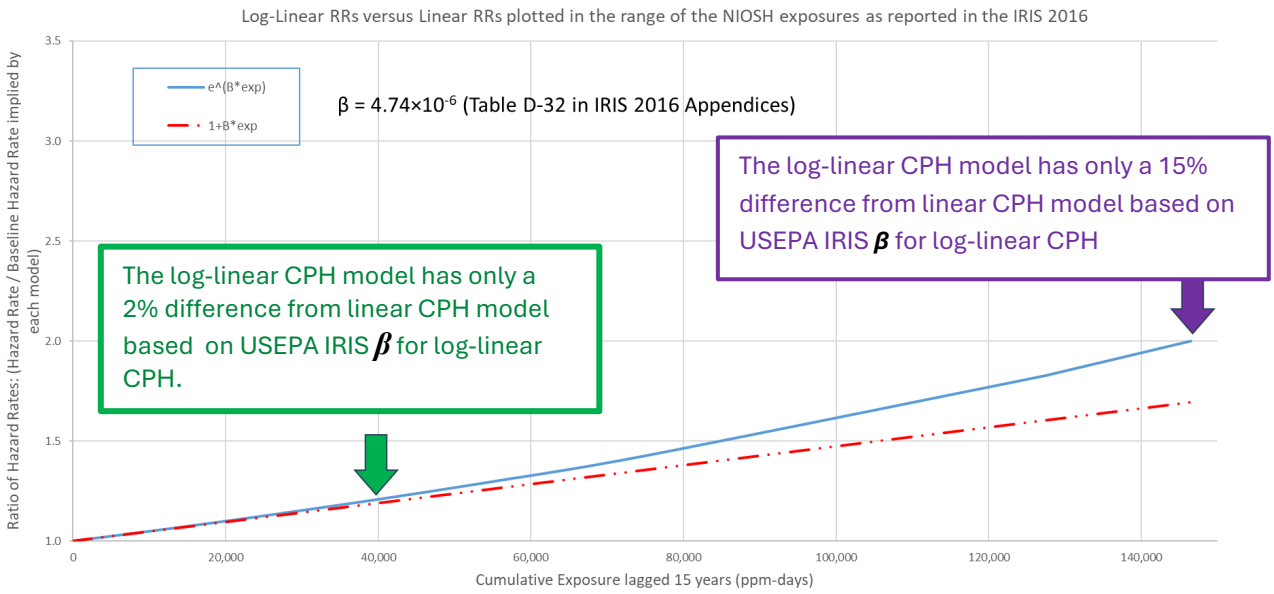
The ACC comments also respond to USEPA's misleading and false claims in USEPA's response to comments (RTC) on the HON and MON rules. Briefly , USEPA IRIS underestimated the p-value for all 2-piece spline models because USEPA failed to account for a's statistical optimization of the knot value (point of inflection) based on maximum or local maximum likelihood. When corrected, there is no difference in p-value or AIC between US EPA's selected model or the TCEQ log-linear CPH model. However, the TCEQ model is more parsimonious (i.e., a simpler model) and, as described in greater detail above, has greater biological plausibility. This is true whether considering lymphoid mortality, breast cancer mortality or breast cancer incidence. It is notable that two different sets of peer reviewers agreed with TCEQ's statistical correction and method of The question

The question raised by US EPA's incorrect analysis is: How much does TCEQ's model approach deviate from a linear model? The comparison can be made by the following steps:

1. Select US EPA IRIS calculated for the log-linear CPH model (US EPA refers to TCEQ Model)
2. Apply to the equation for linear CPH model: $= \mathbf{1} + \boldsymbol{\beta d}$
3. Evaluate (d) for both equations at d = 146,460 ppm-days, which is the highest exposure for a lymphoid case

The results are illustrated in Figure 5

Figure 5. Log-linear RRs versus Linear RRs plotted in the range of the NIOSH exposures as reported in the US EPA IRIS (2016)



These data show that the "TCEQ model" is essentially linear throughout the entire range of cumulative exposure for NIOSH lymphoid cases. US EPA RTC (2022, p. 49) states "It is important to note that both the US EPA and TCEQ's models are linear in the low dose range, a general property recommended for assessing mutagenic carcinogens". In fact, the TCEQ model (log-linear CPH model) is linear throughout the entire range of exposures, consistent with biological evidence. While it's true that the standard log-linear CPH model will eventually bend upwards, it does not "explode" upwards over the range of exposures relevant for risk assessment purposes.

3.3.3. New information on visual fit comparisons along the y-axis

IRIS 2016 correctly notes that relative rate continuous and categorical models cannot be compared meaningfully along the y-axis. Yet, recent US EPA RTC (2022, 2024) falsely equates categorical and continuous models as comparable along the y-axis.

US EPA RTC (2022, p. 52; 20) states incorrectly:

The categorical and continuous results compared utilize the same referent groups – individuals who have no estimated exposure after taking into account the lag period of the modeling. . . Thus, the categorical and continuous relative risk estimates should be in general agreement . . .and may reasonably be compared."

US EPA RTC (2024, p. 101-102, Table 1) then makes an incorrect analysis by numerically comparing continuous model predicted Relative Rate (RRs) at interval midpoints with the categorical model along the y-axis. **This directly contradicts the correct warning in each US EPA IRIS Figure of RRs:**

> "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."

Why is the US EPA IRIS warning note correct?

The CPH model fits data on individual-specific values of hazard rate (HR) in relation to dose "d".

RR (d) = HR (d)/HR (0)

The hazard rate at zero cumulative exposures is HR (0), where HR(0) is the implicit model-specific estimate of the y-intercept. The HR (0) estimate is based on the model-specific equation and all the individual cases (lagged out and unlagged cases).

In contrast, the categorical grouped estimate of HR (0) is less dependent on the individual cases that are not lagged-out. Continuous models with higher implicit baseline risks [HR(0] than for the categorical estimate will appear to be lower than the categorical estimates on graphs of RR.

For example, the linear and log-linear models for the breast cancer incidence (interviewed subcohort) appear lower than the categorical estimates, which are not the individual data modeled. The standard log-linear CPH, the linear CPH and the 2-piece linear spline models are all statistically significant (respective p-values are 0.02, 0.01 and 0.04; see Table 1 of this document). As discussed in detail standard log-linear CPH model has the greatest biological plausibility followed by the linear CPH model.

Consequently, the model selection process used by US EPA IRIS (2016) applied an arbitrary, unprecedented, and meaningless visual-fit approach that was contrary to well-established and widely applied model-selection criteria.

New information regarding the knot and USEPA's steep initial slope

As described above in Section 1, both TCEQ peer reviewers and NAS panel reviewing the TCEQ (2020) agree that the statistical analysis should include the knot as a variable in the model. When corrected, the model fits for the USEPA IRIS 2-linear slope spline model and the USEPA IRIS log-linear CPH model are comparable based on AIC/p-values (Table 1) for Lymphoid Mortality, Breast Cancer Incidence and Breast Cancer Mortality.

The US EPA (2005) cancer risk assessment guidelines note that "a steep slope indicates that errors in an exposure assessment can lead to large errors in estimating risk." These cautions are especially relevant to analysis of NIOSH-cohort cancer mortality and incidence data in light of the high potential for exposure misclassification in the earlier employment years covered by the NIOSH study. As demonstrated above, the absence of any EtO monitoring data prior to 1976, and the paucity of such data prior to 1978, preclude any meaningful validation of the historical NIOSH-cohort exposure model.

Thus, US EPA's own guidance cautions strongly against accepting a steep slope such as with the EtO IRIS value. The only way to reasonably accept it would be to argue for the high reliability of the underlying exposure assessment. However, any such argument would be indefensible. The NIOSH exposure assessment is based on extremely limited exposure data contemporary to the lymphoid and breast cancer cases, and the NIOSH extrapolation used to recover from this data paucity assumes lower, not higher, exposures in earlier period, contrary to everything known about industrial hygiene history.

The basis for the corrections was thoroughly documented and explained by TCEQ (2020, pp. 138-142, Appendix A6.3.1.1 "p-Values and AIC Values"). The TCEQ explanation emphasized that US EPA IRIS (2016, at p. D-38) relied explicitly on p. 12 of a (US EPA SAB, 2015) Science Advisory Board discussion *concerning exploratory sensitivity analyses* of alternative spline-model fits to NIOSH cohort cancer response data which clarified that "The knot is preselected and is not considered a parameter *in these analyses*, consistent with the SAB's concept of parsimony … *the impact of the fixed parameter choices can be evaluated in sensitivity analyses*" [emphasis here added]. However, the US EPA IRIS (2016) assessment clearly *misinterpreted*——and, in US EPA (2022), has similarly continued to misrepresent—its cited SAB recommendation, because the IRIS assessment in fact *did not "fix the knot"* used in its selected lymphoid mortality risk model, *but rather estimated that 1,600-ppm-day knot*, in view of the fact that *the knot's local likelihood value* (conditional on having first rejected the global maximum-likelihood 100-ppm-day knot as reflecting a biologically implausible "step function" response)

had been numerically maximized, as quite clearly explained on p. D-3[19] and as quite clearly plotted in Figure D-14 (at bottom) on p. D-40 and in Figure D-15 on p. D-42 of Appendix D of the IRIS assessment. Consequently, the resulting knot used was in fact an estimated parameter of the model selected for lymphoid cancer mortality risk extrapolation. Thus, the knot is clearly an estimated parameter and needs to be counted in the degrees of freedom.

Before correcting the knot degree of freedom, the p-value for USEPA's lymphoid and breast cancer mortality models were 0.07, or marginally significant. After correction, the p-value=0.14 and 0.15, respectively, is not properly described as statistically significant or marginally significant. The erroneous value of p = 0.07 for lymphoid cancer mortality claimed by US EPA (2016) as noted in Table 1 was repeated more recently in US EPA (2024, Table D-50 at p. 45). For breast cancer incidence, the p-values are 0.04 and 0.02, respectively, for USEPA's and TCEQ's models before and after correction for the knot.

Further, US EPA SAB (2015) did not either direct or suggest that US EPA violate widely applied, standard statistical methods for model selection, such as applying "the Akaike information criterion (AIC) or the Bayesian information criterion (BIC) … approaches [which] correspond to picking the model that maximizes log-likelihood, *subject to a penalty function reflecting the number of model parameters*, thus effectively forcing a trade-off between improving model fit by adding … model parameters versus having a parsimonious description" (NRC 2007 [emphasis added]), as well as in other model-selection contexts (Höge et al., 2018). This trade-off requirement is widely recognized also to pertain specifically to "knots" or inflection points that are estimated in order to fit spline models to data (Berman et al., 1996; Li et al., 2011; Molinari et al., 2001; Rodríguez-Domínguez et al., 2018). USEPA failed to acknowledge its preferred-model knot to be an additional estimated parameter, and then failed to follow well-established statistical methods routinely applied to account properly for the number of estimated model-specific parameters for the purpose of model selection. As a result, US EPA IRIS (2016) violated this standard trade-off requirement in a way that yielded invalid spline-model goodness-of-fit statistics, and consequently an erroneous basis for its selection among models considered.

Based on corrected goodness-of-fit statistics shown in Table 1, the TCEQ standard log-linear CPH and US EPA 2-slope spline models fit comparably well. However, the TCEQ model is simpler (more parsimonious), more consistent with the epidemiological weight-of-evidence, and has much greater biological plausibility compared to the US EPA IRIS 2-slope spline model. This is particularly true in view of complete absence of any published data that corroborate the notably relatively abrupt saturation of EtO-induced DNA or chromosome damage or repair at

[19] At p. D-3: "The shape of the two-piece spline model, in particular the slope of the curve in the low- dose region, depends on the choice of the point of inflection where the two splines are joined. Here, we have chosen the point of inflection based on the best model likelihood, trying a range of points of inflection (knots) across the range of exposure starting from 0 and incrementing by 100, 500, or 1,000 ppm-day intervals. The model likelihood often does not change much across these different points of inflection, but it does change some and we have chosen the point of inflection resulting in the best model likelihood."

subchronic or chronic *in vitro* or *in vivo* EtO concentrations in or around the vicinity of the 2-slope spline model-posited "knot" at 1,600 ppm-days.

3.4. NEW Reality Check of Exposure-Response Models Fit to Epidemiological Data of Workers Exposed to Ethylene Oxide

TCEQ evaluated how well the models considered by IRIS (2016) and TCEQ (2020) predict the actual lymphoid mortality cases reported. This ground-truthing procedure is thoroughly discussed in Appendix 3 "Reality Check of Epidemiological Exposure-Response Model Results for EtO and Lymphoid Cancer Mortality" of TCEQ (2020), including a consideration of a theoretical healthy worker effect. This method is superior to USEPA's incorrect subjective visual fit methods using graphs that are not fit for purpose.

Although the approach by TCEQ (2020) is correct, the following analysis addresses USEPA RTC regarding method of calculating confidence intervals and uses the same intervals reported by Steenland et al. (2004). This updated approach is applied to both breast cancer mortality and lymphoid mortality.

- For calculating the confidence interval, this new analysis uses the exact Poisson confidence intervals (CI) at 90% (i.e., p=0.05) for the expected number of cases for both lymphoid mortality and breast cancer mortality, which is a more conservative stringent requirement than a CI of 95%[20]
  The exposure intervals in the respective publications by Steenland et al. (2004) and Valdez-Flores et al. (2025) were used for the NIOSH and UCC cohorts.

Together with previous comments submitted to USEPA, the analysis presented below address all comments made by USEPA and demonstrates that the standard log-linear CPH model is NOT underpredicting the NIOSH data.

[20] Since this is essentially a one-sided test, the 90% confidence intervals were used to conclude whether a model overpredicts the number of cancer deaths observed at the 5% significance level whenever the lower confidence limit of the estimate exceeds the observed number. This 90th CI is more stringent than 95% CI, in that there is more likelihood of rejection of the model compared to actual number of deaths observed. That is, whenever the probability of observing the actual number of cancer deaths if the model were true (p-value) is less than 0.05.

3.4.1. Ground-truthing of Breast Cancer Mortality Models

IRIS (2016) fit exposure-response models to breast cancer mortality using the NIOSH data followed up through 1998.

Table 7 lists the results of the validation exercise using the quintiles defined by Steenland et al. (2004). The SMR is statistically significantly elevated only at the highest quintile. Consequently, the background model (or linear model with zero slope) statistically significantly underpredicts the observed number of breast cancer deaths (13) in the highest quintile.

In Table 7, IRIS (2016) log-linear and linear models 90% confidence intervals of the predicted number of breast cancer deaths include the observed number of breast cancer deaths in every quintile and overall. IRIS (2016) linear spline model with knot at 700 ppm-days, however, statistically significantly overpredict the observed number of breast cancer deaths in the third quintile and overall.

Table 7. Quintile-specific NIOSH cohort female breast cancer mortalities predicted by Cox log-linear and alternative models with cumulative EtO exposures lagged 20 years

Model	Lagged Out	Quintile 2	Quintile 3	Quintile 4	Quintile 5	Overall
EtO exp (ppm-days)[1]	*0*	(0, 646]	(646, 2779]	(2779, 12,321]	>12,321	*all*
Deaths Observed	*42*	*17*	*15*	*15*	*13*	*102*
Expected	52.06	17.46	14.50	12.68	**6.26**	102.95
SMR 90% Conf. Interval	80.67 (61.35, 104.34)	97.38 (62.04, 146.05)	103.45 (63.76, 159.27)	118.33 (72.93, 182.18)	**207.76** **(122.87, 330.27)**	99.07 (83.51, 116.79)
Expected number of deaths and exact Poisson 90% confidence intervals						
Null RR Model Linear model with zero slope	52.06 (40.73, 65.52)	17.46 (10.83, 25.50)	14.50 (8.46, 21.89)	12.68 (6.92, 19.44)	6.26 (2.61, 11.84)	102.95(85.98, 120.24)
IRIS (2016) Table 4-9 MLE=1.22E-5 (p-value=0.10) Standard Cox model	52.06 (40.73, 65.52)	17.52 (10.83, 25.50)	14.78 (8.46, 21.89)	13.65 (7.69, 20.67)	10.67 (5.43, 16.96)	108.69 (91.49, 126.72)
IRIS (2016) Tab D-25 MLE=2.6779E-5 (p-value=0.07) Linear model	52.06 (40.73, 65.52)	17.60 (10.83, 25.50)	15.11 (9.25, 23.10)	14.72 (8.46, 21.89)	12.28 (6.92, 19.44)	111.77 (94.26, 129.96)
IRIS (2016) Tab D-25 MLE=8.34E-4,-8.07E-4, Knot=700 ppm-day (p-value=0.15) Spline model	52.06 (40.73, 65.52)	21.96 (14.07, 30.24)	23.14 (15.72, 32.59)	21.51 (14.07, 30.24)	14.84 (8.46, 21.89)	133.51 (114.62, 153.59)
IRIS (2016) Tab D-25 MLE=8.30E-4,-8.30E-4, (Zero slope after knot), Knot=700 ppm-day Linear Spline Model	52.06 (40.73, 65.52)	21.96 (14.07, 30.24)	22.85 (14.89, 31.41)	19.96 (12.44, 27.88)	9.83 (4.70, 15.71)	126.66 (108.12, 146.09)

Boldface-statistically significant increased SMR, Green- statistically significant decreased estimated number of cancer deaths, Red- statistically significant increased estimated number of cancer deaths, lack of highlighting indicates model estimates the actual cases.

[1]Actual lag period was 20.05 years and quintile cutoff values for the intervals were 0, 748.76, 2922.0, and 12290.66 ppm-days to match the breast cancer deaths in each quintile reported by Steenland et al. (2004)

3.4.2. Ground-truthing of Lymphoid Cancer Mortality Models

USEPA IRIS (2016) and TCEQ (2020) fit exposure-response models to lymphoid cancer mortality using the NIOSH data followed up through 1998. The lymphoid cancer mortality models reported in USEPA (2016) and TCEQ (2020) (log-linear, linear and 2-piece linear spline models) were evaluated using the lymphoid cancer mortality observed in the NIOSH data through the end of 1998 and the UCC epidemiological data (Valdez-Flores et al. 2025) with follow up through 2013.

Table 8 lists the results of the validation exercise based on the NIOSH data using the quintiles defined by Steenland et al. (2004). The SMR is not statistically significantly elevated overall or at any quintile. Consequently, the background model (or linear model with zero slope) has 90% confidence intervals of the predicted number of lymphoid deaths that includes the number of lymphoid deaths observed in the NIOSH study.

In Table 8, the TCEQ (2020) log-linear model and the IRIS (2016) log-linear and linear models 90% confidence intervals of the predicted number of lymphoid cancer deaths include the observed number of lymphoid cancer deaths in every quintile and overall. IRIS (2016) linear spline models with knot at 1600 ppm-days, however, result in 90% confidence intervals of the predicted number of lymphoid cancer deaths that statistically significantly overpredict the number of lymphoid cancer deaths observed in the NIOSH study overall and in most of the exposure quintiles. The statistically significant overprediction of lymphoid cancer deaths observed in the NIOSH study occurs even after assuming that the risk does not increase (i.e., zero slope) for cumulative exposures above the knot of the IRIS (2016) spline models.

Importantly, Table 8 results indicate that the TCEQ (2020) log-linear model and the IRIS (2016) log-linear and linear models 90% confidence intervals of the predicted number of lymphoid cancer deaths include the observed number of lymphoid cancer deaths in every quintile and overall. In contrast, IRIS (2016) linear spline models with knot at 1600 ppm-days, result in 90% confidence intervals of the predicted number of lymphoid cancer deaths that statistically significantly overpredict the number of lymphoid cancer deaths observed in the NIOSH study overall and in most of the exposure quintiles. The statistically significant overprediction of lymphoid cancer deaths observed in the NIOSH study occurs even after assuming that the risk does not increase (i.e., zero slope) for cumulative exposures above the knot of the IRIS (2016) spline models.

Table 8. Quintile-specific NIOSH cohort male and female lymphoid cancer mortalities predicted by Cox log-linear and alternative models fit to the NIOSH cohort with cumulative EtO exposures lagged 15 years

Model	Lagged Out	Quintile 2	Quintile 3	Quintile 4	Quintile 5	Overall
EtO exp (ppm-days)[1]	0	(0, 1,200]	(1,200, 3,680]	(3,680, 13,500]	>13,500	All
Deaths Observed	*9*	*10*	*11*	*10*	*13*	*53*
Expected	11.48	14.39	7.94	9.13	7.45	50.39
SMR 90% Conf. Interval	78.36 (40.86, 136.71)	70.90 (38.45, 120.23)	166.00 (93.07, 274.71)	109.50 (59.39, 185.69)	143.77 (85.03, 228.55)	105.18 (82.59, 132.22)
Expected number of deaths and exact Poisson 90% confidence intervals						
Null RR Model Linear model with zero slope	11.48 (6.17, 18.21)	14.10 (8.46, 21.89)	6.63 (2.61, 11.84)	9.13 (4.70, 15.71)	9.04 (4.70, 15.71)	50.39 (38.96, 63.29)
TCEQ (2020) Table 8 MLE=2.81E-6 (p-value=0.35) Standard Cox model	11.48 (6.17, 18.21)	14.13 (8.46, 21.89)	6.67 (2.61, 11.84)	9.32 (4.70, 15.71)	10.81 (5.43, 16.96)	52.42 (40.73, 65.52)
IRIS (2016) Tab 4-2 MLE=4.74E-6 (p-value=0.22) Standard Cox Model	11.48 (6.17, 18.21)	14.14 (8.46, 21.89)	6.70 (2.61, 11.84)	9.45 (4.70, 15.71)	12.74 (6.92, 19.44)	54.52 (42.51, 67.74)
IRIS (2016) Tab D-36 MLE=1.227E-5 (p-value=0.13) Linear Model	11.48 (6.17, 18.21)	14.21 (8.46, 21.89)	6.82 (2.61, 11.84)	9.94 (4.70, 15.71)	15.12 (9.25, 23.10)	57.58 (45.18, 71.07)
IRIS (2016) Tab 4-4 MLE=7.58E-4,-7.48E-4, Knot=1600 ppm-day (p-value=0.12) Linear Spline model	11.48 (6.17, 18.21)	20.33 (13.25, 29.06)	14.61 (8.46, 21.89)	20.63 (13.25, 29.06)	24.64 (16.55, 33.75)	91.69 (75.90, 108.32)
IRIS (2016) Tab 4-4 MLE=7.58E-4, -7.58E-4 (Zero slope after knot), Knot=1600 ppm-day (p-value=0.12) Linear Spline model	11.48 (6.17, 18.21)	20.33 (13.25, 29.06)	14.56 (8.46, 21.89)	20.11 (13.25, 29.06)	19.89 (12.44, 27.88)	86.37 (71.34, 102.89)

Boldface-statistically significant increased SMR, Green- statistically significant decreased estimated number of cancer deaths, Red- statistically significant increased estimated number of cancer deaths, lack of highlighting indicates model estimates the actual cases.

[1]Lag period is 15 years and quintile cutoff values for the intervals were 0, 1200, 3680, and 13500 ppm-days to have the same number of lymphoid cancer deaths in each of IRIS (2016) quintile

3.4.3. Ground-truthing of log-linear model for NIOSH lymphoid mortality applied to independent UCC lymphoid mortality

Ground-truthing of the models applied to lymphoid mortality can also be applied to a separate epidemiology Union Carbide Cohort (UCC; Valdez-Flores et al. 2025). Table 9 lists the results of the validation exercise based on the UCC data using the quintiles defined so that every exposure quintile included approximately the same number of lymphoid deaths. The SMR is neither statistically significantly elevated overall nor at any quintile. Consequently, the background model (or linear model with zero slope) has 90% confidence intervals of the predicted number of lymphoid deaths that includes the number of lymphoid deaths observed in the UCC study.

In Table 9, the TCEQ (2020) and IRIS (2016) log-linear models 90 % confidence intervals of the predicted number of lymphoid cancer deaths include the observed number of lymphoid cancer deaths in every quintile and overall. The IRIS (2016) linear model 90% confidence intervals of the predicted number of lymphoid cancer deaths include the observed number of lymphoid cancer deaths in the first four quintiles and overall. IRIS (2016) linear spline models with knot at 1600 ppm-days, however, result in 90% confidence intervals of the predicted number of lymphoid cancer deaths that statistically significantly overpredict the number of lymphoid cancer deaths observed in the UCC study overall and in the higher exposure quintiles. The statistically significant overprediction of lymphoid cancer deaths observed in the UCC study occurs even after assuming that the risk does not increase for cumulative exposures above the knot of the IRIS (2016) spline models.

Table 3. Quintile-specific UCC cohort male and female lymphoid cancer mortalities predicted by Cox log-linear and alternative models fit to the NIOSH cohort with cumulative EtO exposures lagged 15 years

Model	Lagged Out	Quintile 2	Quintile 3	Quintile 4	Quintile 5	Overall
EtO exp (ppm-days)[1]	0	(0, 3,000]	(3,000, 7,150]	(7,150, 20,000]	>20,000	All
Deaths Observed	*0*	*6*	*6*	*6*	*7*	*25*
Expected	1.92	6.26	4.51	5.08	8.44	26.20
SMR 90% Conf. Interval	0.00 (0.00, 154.64)	95.92 (41.72, 189.23)	133.03 (57.86, 262.44)	118.10 (51.36, 232.98)	82.98 (38.91, 155.80)	95.42 (66.34, 133.25)
Expected number of deaths and exact Poisson 90% confidence intervals						
Null RR Model Linear model with zero slope	1.92[2] (0.05, 4.74)	6.26 (2.61, 11.84)	4.51 (1.37, 9.15)	5.08 (1.97, 10.51)	8.44 (3.98, 14.43)	26.20 (18.22, 36.08)
TCEQ (2020) Table 8 MLE=2.81E-6 (p-value=0.35) Standard Cox model	1.92 (0.05, 4.74)	6.28 (2.61, 11.84)	4.58 (1.37, 9.15)	5.26 (1.97, 10.51)	10.06 (5.43, 16.96)	28.09 (19.90, 38.39)
IRIS (2016) Tab 4-2 MLE=4.74E-6 (p-value=0.22) Standard Cox Model	1.92 (0.05, 4.74)	6.29 (2.61, 11.84)	4.62 (1.37, 9.15)	5.39 (1.97, 10.51)	11.48 (6.17, 18.21)	29.70 (20.75, 39.54)
IRIS (2016) Tab D-36 MLE=1.227E-5 (p-value=0.13) Linear Model	1.92 (0.05, 4.74)	6.34 (2.61, 11.84)	4.80 (1.37, 9.15)	5.86 (1.97, 10.51)	14.53 (8.46, 21.89)	33.45 (24.15, 44.13)
IRIS (2016) Tab 4-4 MLE=7.58E-4,-7.48E-4, Knot=1600 ppm-day (p-value=0.12) Linear Spline model	1.92 (0.05, 4.74)	10.73 (5.43, 16.96)	10.02 (5.43, 16.96)	11.65 (6.17, 18.21)	23.13 (15.72, 32.59)	57.44 (45.18, 71.07)
IRIS (2016) Tab 4-4 MLE=7.58E-4, -7.58E-4 (i.e., Zero slope after the knot), Knot=1600 ppm-day (p-value=0.12) Linear Spline model	1.92 (0.05, 4.74)	10.71 (5.43, 16.96)	9.86 (4.70, 15.71)	11.10 (6.17, 18.21)	18.41 (11.63, 26.69)	52.01 (40.73, 65.52)

Table Notes:
Boldface-statistically significant increased SMR, Green- statistically significant decreased estimated number of cancer deaths, Red- statistically significant increased estimated number of cancer deaths, lack of highlighting indicates model estimates the actual cases.

[1]Lag period is 15 years and quintile cutoff values for the intervals were 0, 1550, 4500, and 18000 ppm-days to have the same number of lymphoid cancer deaths in each quintile
[2]The 90% confidence interval of the expected number of lymphoid cancer deaths indicates that the predicted number of lymphoid cancer deaths in the DOW study statistically significantly overestimates the observed number of lymphoid cancer deaths (zero) for all models. This is due to the fact that there were zero lymphoid cancer deaths in the lagged out group and should not be interpreted as a statistically significant overestimation but rather a characteristic of the bunding procedure (exact Poisson confidence intervals).
Note: Based on IRIS Table 4-4, the slope after the knot for the model is -7.48E-04. The assumption of perfect negative correlation is consistent with the covariance values obtained by US EPA for linear two-piece spline model; e.g., see footnote to Table D-36 in the appendices of US EPA's report.

3.4.4. Summary

Exposure-response models fit to mortality endpoints in epidemiological data were evaluated to determine the validity of their predicted risks. The standard Cox (log-linear) model best predicted the observed number of cancer deaths for both endpoints. The linear model fit to all the data was, for the most part, also validated. However, spline models consistently overpredicted the observed number of cancer death for both endpoints and across different studies.

4.0. USEPA Question #5: Is there new information related to human exposure to EtO, including information on occupational, smoking, background or endogenous exposures that may be relevant to estimating the dose-response relationship of EtO carcinogenicity?

4.1 Overview

Ethylene oxide (EtO) is a chemical used to sterilize medical equipment and the synthesis of other chemicals. It is also classified as a cancer-causing substance and for this reason exposures to EtO are of great interest to regulators and the general public. In addition to exogenous exposures to EtO in air, the human body naturally produces EtO through normal biological processes—this is called endogenous production. Endogenous exposures to EtO and their relative contribution to total exposures in the U.S. population have been well characterized in the scientific literature published since US EPA's 2016 IRIS Assessment (US EPA, 2016). Key points from this body of literature are summarized below:

- A majority (>90%) of a nonsmoker's total EtO exposure comes from the body's own internal production, rather than from the air or nearby factories. Endogenous exposure estimates are based on high quality data sets, including national blood testing surveys (CDC NHANES) conducted across the U.S. population, and US EPA's air monitoring data for EtO across the U.S.
- Despite their reliance on high quality data sets, estimates of endogenous exposure have been repeatedly criticized by US EPA in response to comments and in the published literature.
    - Initial estimates of endogenous exposure by Kirman and Hays (2017) were initially criticized by US EPA for relying upon pooled data for unexposed populations, over concerns these populations may have EtO exposures which would serve to result in overestimates for endogenous exposure.
        - With the subsequent release of EtO biomonitoring data from CDC NHANES, updated analyses (Kirman et al., 2025) revealed that the pooled data used in the initial study are in excellent agreement with exposures measured in U.S. nonsmokers (see Figure 2 of Kirman et al. 2025).
    - US EPA criticized the reliance upon data for highly exposed workers to characterize a linear relationship between EtO concentrations in air and measured biomarkers as "*not factually substantiated*" for its application to low background exposures to EtO (US EPA, 2022).
        - Subsequent analysis of three lines of evidence: (1) NHANES data for smokers; (2) physiologically based pharmacokinetic (PBPK) model predictions; and (3) pharmacokinetic model predictions; all support the linear

relationship used in estimating endogenous exposures to EtO extends to low background exposures, and therefore is considered to be factually substantiated (Kirman et al., 2025).

- A recent study by US EPA scientists (Lin et al., 2025) examined the published literature for ethylene in exhaled breath and conducted PBPK modeling, from which they conclude that endogenous exposures to EtO are much lower than previously estimated. The authors estimated this pathway accounts for less than 20% of the EtO-related markers in blood, leaving over 80% unexplained.
    - A critical review of that US EPA's assessment identified important limitations: (1) overly narrow consideration of the ethylene exhaled breath data and failure to consider the importance of analytical methods; and (2) failure to consider the lack of validation of the PBPK model used to predict background biomarker concentrations, in which it substantially underpredicts measurements across mice, rats, and humans (predicting only 5–26% of actual levels). This underprediction problem may be due to the model not accounting for three known biological pathways that produce ethylene and EtO in the body (Kirman and Bus, 2026).

- The full body of research on endogenous EtO supports a key conclusion: risk assessment and risk management for EtO should account for the body's own natural production of this chemical. Regulatory decisions that ignore this endogenous context do not reflect best available science or Gold Standard Science.
- Characterization of endogenous exposures serves as a reality check that can be used to put risk-based exposure levels into context when making risk management decisions for EtO:
    - Researchers developed two tools to put EtO exposure in context: the Endogenous Equivalent (EE), which translates the body's internal EtO production into an equivalent air concentration, and the Total Equivalent (TE), which combines internal and external exposure into a single number for comparison (Kirman and Hays, 2017; Kirman et al., 2021, 2025).
    - When the TE framework was applied to communities near EtO-emitting industrial facilities, the facilities' contributions to residents' total EtO exposure were very small—typically less than 1% of what people already experience from their own body's natural production (Sheehan et al., 2021).
    - A case study of a sterilization facility in Covington, Georgia confirmed these findings using both air monitoring data and atmospheric dispersion modeling: the facility did not meaningfully increase local EtO levels above normal background (Lewis et al., 2022).

- There is a ~260-fold difference between the risk-based air concentration limits set by the U.S. USEPA and those set by the Texas Commission on Environmental Quality (TCEQ). The US EPA's risk-based limits are so low that they fall below what instruments can even detect and far below what the body naturally produces, making them difficult to use in practice. In contrast, the TCEQ's limits fall within the range of normal background exposure and can be used to meaningfully distinguish whether a facility is adding significant EtO to nearby communities.

4.1.1 Derivation of Endogenous Equivalent Values to Support Risk Assessment and Risk Management Decisions for an Endogenous Carcinogen: Ethylene Oxide. Kirman CR, Hays SM. Regulatory Toxicology and Pharmacology, 2017; 91:165–172

- Objective: To introduce and derive endogenous equivalent (EE) values for EtO—defined as the air concentration of EtO that would produce exposure equivalent to that arising from endogenous production—in order to provide context for risk assessment and risk management decisions.
- Methods: The authors conducted a meta-analysis of published hemoglobin N-(2-hydroxyethyl)valine (HEV) adduct data from populations without known occupational or environmental EtO exposure. A linear relationship between HEV adducts and continuous EtO air concentration (HEV [pmol/g] = 10.9 × [EtO, ppb]) was applied to convert background HEV levels into equivalent air concentrations, yielding a distribution of EE values.
- Results: Background HEV adduct levels in unexposed nonsmoking populations ranged from approximately 14 to 75 pmol/g globin. These corresponded to EE values ranging from 0.13 to 6.9 ppb (median ~2.1 ppb). This demonstrates that endogenous EtO production in the general population produces internal exposure substantially exceeding typical ambient air concentrations. The EE framework provides a quantitative benchmark for comparing exogenous exposures against the body's own EtO burden.
- **Conclusions: The EE concept offers a scientifically grounded tool for placing exogenous EtO exposures into biological context. Because endogenous production dominates total EtO exposure in nonsmokers, risk assessments and risk management decisions that do not account for endogenous exposure may lack practical relevance. The authors recommend incorporating EE values into regulatory frameworks to improve the scientific basis for decision-making.**

4.1.2 Ethylene Oxide Review: Characterization of Total Exposure via Endogenous and Exogenous Pathways and Their Implications to Risk Assessment and Risk Management.

Kirman CR, Li AA, Sheehan PJ, Bus JS, Lewis RC, Hays SM. Journal of Toxicology and Environmental Health, Part B: Critical Reviews, 2021; 24(1):1–29

- Objective: To provide a comprehensive review of all known endogenous and exogenous sources of EtO exposure in the general population, update the endogenous equivalent (EE) framework with newly available CDC NHANES biomonitoring data, and evaluate the implications for risk assessment and risk management.
- Methods: The authors compiled and analyzed data from three CDC NHANES sampling periods (2013–2016) for HEV adducts across the U.S. population. They characterized endogenous pathways (gastrointestinal bacterial production, systemic enzymatic production via lipid peroxidation and amino acid metabolism) and exogenous pathways (ambient air, tobacco smoke, industrial releases). The HEV-to-air-concentration relationship was refined (11.3 ± 4.2 pmol/g per ppb), and updated EE distributions were derived.
- Results: Approximately 93% of total EtO exposure in nonsmokers is endogenous in origin. The updated EE distribution based on NHANES data yielded a central tendency of approximately 1.9 ppb with a 95th percentile near 4.0 ppb. Cigarette smoking contributes meaningfully to exogenous exposure, while ambient air EtO (typically <0.2 ppb) and near-facility industrial contributions are minor relative to total body burden. The review demonstrated that US EPA risk-specific concentrations (RSCs) of 0.00011–0.011 ppb are orders of magnitude below endogenous background and ambient detection limits, raising questions about their practical utility.
- **Conclusions: The dominance of endogenous EtO exposure means that risk management decisions should account for endogenous background when interpreting ambient monitoring data. US EPA RSCs fall well below measurable and background levels, limiting their utility as practical risk management tools. Risk-based concentrations that incorporate endogenous context, such as those derived by TCEQ, are better positioned to inform regulatory decisions.**

Estimates of endogenous exposure to EtO continue to be questioned by US EPA (US EPA, 2022). Specific criticisms of the approach to calculating TE and EE values have included (1) a position that that the proposed linear relationship between HEV biomarkers and exposures to workers has not been factually substantiated for its application to low-level and background exposures to EtO; (2) the EE analysis relies on the measurement of EtO HEV adducts and does not include direct measurements of endogenous EtO levels; (3) since cancer potency estimates for EtO are expressed in terms of extra risk (i.e., above background), endogenous exposures are not included in the assessment; (4) other biomarker data (e.g., ET in exhaled breath) are not

consistent with a large source of endogenous exposure; and (5) background ambient EtO measurements away from emitting facilities are unreliable.

4.1.3 Ethylene Oxide Exposure in U.S. Populations Residing Near Sterilization and Other Industrial Facilities: Context Based on Endogenous and Total Equivalent Concentrations. Sheehan PJ, Lewis RC, Kirman CR, Watson HN, Winegar ED, Bus JS. International Journal of Environmental Research and Public Health, 2021; 18:18020607

- Objective: To introduce the total equivalent (TE) concentration metric—combining endogenous and exogenous EtO exposure into a single framework—and evaluate modeled ambient EtO concentrations near eight EtO-emitting industrial facilities in the context of total population exposure.
- Methods: The authors developed the TE metric by summing endogenous equivalent concentrations (derived from NHANES HEV data) with exogenous ambient EtO concentrations from US EPA National Air Toxics Assessment (NATA) and facility-specific modeling. Eight EtO-emitting facilities across the U.S. were evaluated. Facility contributions were assessed relative to population TE distributions derived from NHANES biomonitoring data.
- Results: The TE framework demonstrated that near-facility ambient EtO concentrations contribute a small fraction of total population exposure. None of the eight evaluated facilities produced TE concentrations exceeding the 95th percentile of the general population TE distribution. The contribution of individual facility emissions to nearby residents' total EtO burden was typically less than 1% of the population TE range. This approach provided a population-health perspective that risk-based concentrations alone cannot convey.
- **Conclusions: The TE framework offers regulators a practical tool for contextualizing facility-specific EtO exposures within overall population exposure. Because endogenous production dominates total EtO body burden, facility emissions—even at the modeled upper bounds—do not meaningfully shift exposed populations beyond the range of normal background variability. This context is essential for proportionate and science-based risk management.**

4.1.4. Assessment of Ethylene Oxide Exposures from Ambient Air Monitoring Near a Sterilization Facility in Covington, Georgia. Lewis RC, Sheehan PJ, Kirman CR, Watson HN, Winegar ED, Bus JS. International Journal of Environmental Research and Public Health, 2022; 19(16):9984

- Objective: To assess ambient EtO concentrations near a sterilization facility in Covington, Georgia using both USEPA monitoring data and AERMOD dispersion modeling, and to evaluate the facility's contribution to local residents' total EtO exposure in the context of the TE framework.
- Methods: The authors analyzed ambient EtO monitoring data collected by US EPA near the Covington, Georgia facility. AERMOD atmospheric dispersion modeling was used to estimate the facility's contribution to local ambient EtO. Monitored and modeled concentrations were compared against background EtO levels measured at upwind and regional monitoring stations. The TE framework from Sheehan et al. (2021) was applied to place facility contributions in the context of total population exposure.
- Results: Ambient monitoring data revealed that EtO concentrations at the Covington site were largely consistent with regional background levels. AERMOD modeling confirmed that the facility's maximum contribution to ambient EtO was small relative to background. When placed in the TE framework, the facility's contribution to total resident exposure was negligible—well within the normal range of population variability in endogenous EtO production. Background EtO concentrations measured at upwind sites were similar to downwind measurements, further supporting the conclusion that the facility was not a meaningful driver of local EtO levels.
- Conclusions: Monitored ambient EtO near the Covington facility does not indicate elevated community exposure attributable to the facility. Both measured and modeled data confirm that facility emissions contribute negligibly to residents' total EtO body burden. This case study reinforces the importance of interpreting ambient monitoring data in the context of endogenous background and total equivalent concentrations rather than relying solely on risk-based benchmarks.

4.1.5. Characterization of Background Exposures to Ethylene Oxide in the United States: A Reality Check on Theoretical Health Risks for Potentially Exposed Populations near Industrial Sources. Kirman CR, Sheehan PJ, Li AA, Bus JS, Su SH, Dopart PJ, Watson HN, Moynihan EE, Reiss R. International Journal of Environmental Research and Public Health, 2025; 22(4):597

- Objective: To provide a comprehensive, updated characterization of background EtO exposure in the U.S. population using bias-corrected NHANES data, Monte Carlo uncertainty/variability analysis, and validation against smoker data; to address specific criticisms made by US EPA on endogenous EtO estimates; and to evaluate whether US EPA and TCEQ risk-specific concentrations (RSCs) have practical utility as risk management tools.
- Methods: The authors utilized bias-corrected HEV adduct data from three CDC NHANES sampling periods (2013–2020), applied Monte Carlo simulation to propagate uncertainty and variability in TE and EE estimates, validated the HEV-to-EtO relationship using smoker data as an internal reference, and incorporated ambient monitoring data from Georgia and Utah sterilization facilities. US EPA and TCEQ RSCs were compared against population TE distributions.
- Results: The updated analysis confirmed that endogenous EtO production accounts for approximately 93% of total nonsmoker exposure (central tendency EE ~1.9 ppb; 95th percentile ~4.0 ppb). The linear HEV-to-air-concentration relationship was validated from occupational through smoking to ambient exposure levels. US EPA RSCs (0.00011–0.011 ppb) fall below ambient detection limits and well below endogenous background, rendering them impractical for risk management. By contrast, TCEQ RSCs (0.24–24 ppb) align with the population TE range and can meaningfully distinguish facility-related exposure increments. Near-facility ambient EtO contributes less than 1% of total population exposure. The 260-fold difference between US EPA and TCEQ RSCs arises from differing dose–response model assumptions.
- Conclusions: US EPA RSCs lack practical utility because they are below detectable and background levels, making them unsuitable for informing risk management actions. TCEQ RSCs, which are consistent with population TE distributions, offer a scientifically defensible and actionable framework. The analysis further supports the conclusion that industrial EtO emissions contribute negligibly to total population exposure. The authors recommend that regulators adopt risk management frameworks that incorporate endogenous context and validated biomonitoring data.

4.1.6. Uncovering the Connection: Ethylene Exposure and Endogenous Ethylene Oxide Levels in Humans. Lin Y-S, Thayer KA, White P, Morozov V, Persad AS. Journal of Exposure Science & Environmental Epidemiology, 2025 (Published online November 20, 2025)

- Objective: To investigate the impact of total ethylene (ET) exposure—both external (ambient air) and internal (endogenous production)—on endogenous EtO levels in non-smoking, non-occupationally exposed populations, using literature review and pharmacokinetic modeling.
- Methods: The authors conducted two systematic literature searches (January 1960–April 2023) to identify studies measuring exhaled ET in breath and HEV adducts from ET exposure. From 1,179 unique records, 28 studies met inclusion criteria. Two pharmacokinetic approaches were employed: (1) application of the 2018 Filser and Klein PBPK model incorporating endogenous ET sources, and (2) a simplified algebraic kinetic model characterizing the relationship between exhaled ET and endogenous EtO kinetics. Study quality was assessed using a confidence-rating framework for breath ET measurement methodologies.
- Key Findings: Exhaled ET concentrations varied widely across studies (median 12 ppb; range 0.2–740 ppb). High-confidence studies using purified air showed lower ET concentrations (0.2–1.2 ppb), suggesting limited internal ET contribution. The PBPK model predicted that in an ET-free environment, less than 1–2 ppb exhaled ET can arise from internal sources. With nominal ambient ET exposure of 15 ppb, predicted exhaled ET reached 16 ppb. Less than 3.5 pmol HEV/g globin—under 20% of the approximately 20 pmol/g background HEV—could be attributed to ET-derived EtO. The authors concluded that endogenous EtO from ET exposure appears much lower than previously suggested in the literature.
- Conclusions: The analysis indicates that ET-to-EtO conversion contributes less to background HEV adducts than earlier studies implied, with over 80% of background HEV remaining unexplained by ET-related pathways alone. The authors note that neither internal ET production nor typical ambient ET exposure significantly impacts circulating EtO levels. Given Eto's carcinogenic potential, the authors call for further research to identify the sources of background HEV and to refine PBPK models and analytical methodologies.

4.1.7. Regarding Ethylene Exposure and Endogenous Ethylene Oxide Levels in Humans (Lin et al., 2025). Kirman CR, Bus JS. Submitted to Journal of Exposure Science & Environmental Epidemiology, 2026

- Objective: To critically evaluate the data, assumptions, and PBPK modeling approach used by Lin et al. (2025), with particular focus on analytical method variability, ambient ET exposure assumptions, model validity for predicting background HEV, and the biological adequacy of endogenous ET pathways represented in the model.
- Methods: The authors conducted a critical review examining four key areas: (1) variability between breath ET analytical methods—laser-based photoacoustic spectrometry (LBP) versus gas chromatography (GC); (2) appropriateness of ambient air ET exposure assumptions relative to population monitoring data; (3) PBPK model validity for predicting HEV across species (mice, rats, humans); and (4) completeness of the endogenous ET production pathways included in the Filser and Klein PBPK model.
- Results: LBP-measured ET was far more variable (0.2–490 ppb) than GC-measured ET (6.8–38.2 ppb), with LBP values previously shown to be inconsistent with known ET toxicokinetics. The ambient ET exposures assumed by Lin et al. (15 and 100 ppb) represent high-end rather than typical levels based on Health Canada monitoring data (50th percentile: 3.1 ppb; 95th percentile: 13.6 ppb). The PBPK model predicted only 5–26% of directly measured HEV across mice, rats, and humans, indicating systematic underprediction. Inspection of the model code revealed that endogenous ET is represented as a single pathway in the arterial compartment, an oversimplification given that the literature supports at least three distinct pathways: GI bacterial production, hepatic metabolism of dietary components, and oxidative stress-induced production.
- Conclusions: Exogenous air exposures and proposed artifactual sources cannot account for observed background HEV levels. The underrepresentation of endogenous ET production pathways in the current PBPK model—particularly gut microbiome activity and hepatic metabolism of dietary components—is the most plausible explanation for the model's systematic HEV underprediction. PBPK model refinement to include all three known endogenous ET pathways is needed for reliable source apportionment. Accurate attribution of HEV to endogenous versus exogenous sources is critical for placing risk-based concentrations into proper context for regulatory decision-making.

4.1.8. Lin, Y.-S., Morozov, V., & Wu, K.-Y. (2026). Integrating PBPK modeling and tobacco biomarkers to interpret N-(2-hydroxyethyl)valine (HEV) hemoglobin adducts in the U.S. population. International Journal of Hygiene and Environmental Health, 272, 114721.

- Objective: EtO is a recognized human carcinogen, and cigarette smoke is the dominant environmental source of exposure in the general population. EtO alkylates the N-terminal valine of hemoglobin to form the adduct HEV, which has a ~120-day lifespan in red blood cells and is therefore widely used as a long-term biomarker of EtO exposure. However, HEV levels vary widely even among smokers reporting similar cigarette consumption, and low but measurable HEV is also found in non-smokers, reflecting endogenous production, passive smoke, and other environmental sources. The authors set out to quantify how much of the U.S. population's HEV burden is attributable to active smoking and to evaluate whether combining HEV with nicotine-derived and other tobacco biomarkers improves exposure assessment.
- Methods: Using NHANES 2013–2014 data (n = 1,770 adults with HEV plus supplemental smoking biomarkers), participants were classified as active smokers, passive smokers, or non-smokers based on serum cotinine. Self-reported cigarettes-per-day and cotinine-derived estimates of EtO intake were entered into the 2018 Filser and Klein physiologically based pharmacokinetic (PBPK) model, which accounts for body-weight-specific distribution and metabolism, to predict HEV. Predicted values were compared to measured HEV and to other tobacco biomarkers — serum cotinine, urinary total cotinine, NNAL (4-methylnitrosamino-1-3-pyridyl-1-butanol), and 2-HEMA (2-hydroxyethyl mercapturic acid). Survey-weighted regression was used to examine demographic, behavioral, and health predictors of HEV variability.
- Results: Measured HEV was highest in active smokers (median 186 pmol/g globin), intermediate in passive smokers (29 pmol/g), and lowest in non-smokers (28 pmol/g). PBPK-predicted HEV matched observed levels well in active smokers (median 166 pmol/g) but was consistently 20–30 pmol/g globin lower than measured values in passive smokers and non-smokers, suggesting that non-smoking sources (endogenous formation, ambient ethylene, other environmental exposures) are not captured by cigarette-derived exposure inputs. Body weight was an important determinant: heavier individuals had lower HEV for the same consumption, consistent with dilution across a larger blood/hemoglobin pool. In multivariable regression, smoking status was the strongest predictor, with BMI and poverty-to-income ratio also important; age, sex, and race/ethnicity showed smaller effects. Among active smokers, HEV correlated strongly with serum cotinine (Spearman up to 0.94), moderately with NNAL and 2-HEMA, but correlations were weaker in passive

smokers and non-smokers, reflecting the dominance of background EtO sources at low exposures.

- **Conclusions: Cigarette smoking is the primary source of HEV in the U.S. population, but HEV alone cannot distinguish tobacco exposure from endogenous or ambient background at low intake levels. Combining PBPK modeling with multiple tobacco biomarkers (cotinine, NNAL, 2-HEMA) provides a more complete picture, separates short-term from long-term exposure, and strengthens exposure assessment for population health research and tobacco control. Limitations include the cross-sectional NHANES design, uniform metabolism assumptions in the PBPK model, variable cigarette yields, and absence of genetic (e.g., GSTT1) and co-exposure data.**

4.2. Conclusions on Body of Publications Addressing the Question: *How should endogenous EtO production be accounted for when evaluating health risks from exogenous environmental exposures?*

These publications, all published since US EPA's 2016 IRIS assessment (US EPA, 2016) form a coherent body of work addressing a central question in EtO risk assessment: *how should endogenous EtO production be accounted for when evaluating health risks from exogenous environmental exposures?* Collectively, they establish the scientific foundation for the endogenous equivalent (EE) and total equivalent (TE) concentration frameworks and apply these frameworks to real-world regulatory scenarios. Kirman and Hays (2017) introduced the EE concept and demonstrated through meta-analysis that endogenous EtO production generates internal exposures far exceeding typical ambient air concentrations. Kirman et al. (2021) substantially expanded this foundation by integrating CDC NHANES biomonitoring data, characterizing all known endogenous and exogenous EtO pathways, and establishing that approximately 93% of total EtO exposure in nonsmokers is endogenous in origin. These two papers provide the quantitative underpinning for all subsequent analyses.

This framework has been applied to risk assessment contexts for EtO facilities. Sheehan et al. (2021) operationalized these findings by developing the TE concentration metric, demonstrating across eight industrial facilities that near-facility ambient EtO contributes negligibly to total population exposure. Lewis et al. (2022) applied the same framework to a specific case study in Covington, Georgia, using both monitored and AERMOD-modeled data to confirm that the local sterilization facility did not meaningfully elevate community EtO exposure above background. Kirman et al. (2025) provided the most comprehensive update, incorporating bias-corrected NHANES data, Monte Carlo simulation, smoker validation, and a direct comparison showing that US EPA risk-specific concentrations (RSCs) fall below detection and background levels, while TCEQ RSCs align with the population TE range and offer practical risk management utility.

US EPA has continued its criticism of endogenous EtO estimates. Most recently, Lin et al. (2025) relied upon a review the literature for ET in exhaled breath along with PBPK modeling to

suggest that ET-derived EtO contributes less than 20% of background HEV adducts, with over 80% remaining unexplained. Kirman and Bus (2025) provided a critical response, identifying methodological limitations in the LBP breath data used, the high-end ambient ET assumptions, and a systematic underprediction of HEV by the PBPK model across species. This underprediction by the PBPK model may be due to an oversimplified representation of endogenous ET production and recommended that all three known endogenous pathways (GI microbial, hepatic dietary metabolism, oxidative stress) be incorporated into future modeling efforts.

Overall, several conclusions are supported by this body of publications. First, endogenous EtO production dominates total exposure in the general nonsmoking population, a finding consistently supported by available biomonitoring and monitoring data for EtO. Second, near-facility industrial EtO emissions contribute negligibly to total population exposure when placed in the context of endogenous background. Third, the large discrepancy between US EPA and TCEQ risk-specific concentrations (approximately 260-fold) has profound practical implications: US EPA RSCs fall below measurable and background levels, while TCEQ RSCs provide actionable benchmarks. Fourth, the PBPK models currently used to predict HEV from ET exposure systematically underpredict measured values, likely due to incomplete representation of endogenous ET pathways, underscoring the need for model refinement before such tools are used in regulatory decision-making. Together, these papers advocate for risk assessment frameworks that integrate endogenous context, validated biomonitoring data, and scientifically defensible dose–response modeling to risk assessment and risk management decisions based on best available science.

5. Question #6: Is there <u>additional</u> information on the comments of the 2016 IRIS value provided by commenters in the context of the more recent proposed CMAS NESHAP (January 22, 2025) (see Docket ID USEPA–HQ–OAR–2024–0303, comment)numbers 0060, 0061, 0068, 0076.

The American Chemistry Council submitted comments on the proposed CMAS NESHAP on April 14, 2026 that included an Attachment 1 regarding questions raised by ACC that have not been addressed by USEPA RTC on the MON or HON. EOSA raises these questions again to USEPA.

5.1. Biological Plausibility as an Essential Element to Informing the Shape of the Dose-Response Model

It is important that USEPA address the following questions as essential element to inform the biological plausibility for the selection of the dose-response model:

1. **Does USEPA agree that, given the identical PBPK-model-supported toxicokinetics of EtO in animals and humans, data from animal studies are predictive of the shape of the EtO dose-response in humans?**
2. **If not, what evidence does USEPA have to depart from USEPA's default assumption that animals are predictive of humans?**
3. **Does USEPA agree that the toxicokinetic and most relevant biological data do not indicate a steep dose response at lower exposures? If not, why not?**
4. **Does USEPA agree that there is no evidence of a steep exposure response at lower EtO equivalent exposures in rats?**

The following summarizes our response to these questions on biological plausibility and why they are important for USEPA to address.

The USEPA has not included EtO-specific data or animal/human modes of action for reactive epoxides in dose-response selection. The 2022 RTC overlooked integrating biological evidence, while EOSA contends that both biological and epidemiological findings support the standard CPH model rather than a steep low-exposure response. The 2024 HON RTC did not address ACC's 2023 critique of earlier USEPA analysis.

The USEPA SAB emphasized that selecting an EtO cancer risk model must be anchored on the biological plausibility of its dose-response shape. EtO offers substantial biological evidence—including molecular, toxicokinetic, genotoxicity, carcinogenicity, and mode-of-action studies—relevant to assessing this plausibility. Yet, USEPA IRIS (2016) has not sufficiently integrated these data to support its favored 2-piece linear spline model, which features a steep initial slope at low exposure and flattens at higher EtO levels found in occupational settings. USEPA's

argument rejects other models as implausible but does not adequately justify biological plausibility for its chosen model. This creates an artificially limited choice for plausibility.

Simply put, the biological plausibility of the USEPA-selected dose-response model is in reality driven mostly by incorrect model statistics and misinterpretation of visual fitting of the epidemiological dose-response data and is additionally largely absent an USEPA SAB-demanded examination of the strengths and weaknesses of foundational biological plausibility assessments necessary to the selection of a biologically-defensible human dose-response model option.

EOSA notes that ACC (2020; 2023) provided detailed comments that argued that integrating multiple lines of biological and toxicological evidence supports the TCEQ single-linear dose-response model's plausibility over the USEPA's 2-piece linear slope spline model. The USEPA RTC (2024) rejected this, stating insufficient knowledge about the mode of action to use integrated evidence for dose-response modeling, and relied instead on human data for EtO's carcinogenicity. This reflects USEPA's admission that it has not met SAB's requirement for biological plausibility in its selected dose-response model. USEPA RTC (2022, 2024) also presented non-substantive toxicological counter-analyses challenging ACC's position on the adequacy of integrated biological evidence for the TCEQ model.

The USEPA RTC (2022, 2024) presents varied dose-response patterns from EtO genotoxicity and animal carcinogenicity studies, highlighting inconsistency among supralinear, linear, and sublinear responses. It finds these insufficient to reliably inform plausible cancer dose-response models based on NIOSH epidemiology data. The USEPA RTC (2022) concludes that high-dose rodent testing is not predictive of human dose-response patterns.

Kirman et al. (2021) challenged the above conclusions by highlighting a combined apical tumorigenic dose-response pattern in rats exposed to EtO and ethylene (using PBPK modeling; Filser and Klein, 2018). This approach considers all mechanisms contributing to tumor formation, including EtO detoxification, DNA adducts, genotoxicity, and cytotoxicity. USEPA and ACC agree that EtO's weak genotoxicity supports a conservative linear dose-response assumption, though high-dose non-genotoxic effects cannot be ruled out. The rat dose-response range reflects exposures seen in the NIOSH cohort and does not show the steep-to-shallow response predicted by USEPA's two-linear spline model.

PBPK models show that EtO blood concentrations in humans, rats, and mice are linear and overlap for exposures up to 6 hr/day at 200 ppm. At higher levels, mouse blood EtO rises disproportionately due to reduced glutathione detoxification from EtO-induced depletion. This suggests similar EtO toxicokinetics across species between 1 and 200 ppm, meaning toxicokinetics are unlikely to explain cross-species differences in sensitivity or tumorigenicity. Notably, human EtO blood concentrations do not show dose-dependent steep increases followed by shallowing at exposures above 1 ppm.

Cancer bioassays conducted in rats show that high EtO concentrations (10 ppm - 100 ppm) significantly increase cancer incidences, while similar tests for ethylene at high concentrations (300–3,000 ppm) do not produce increased tumor rates. HEV measurements indicate that elevated ethylene exposure leads to much higher systemic levels of EtO (up to 160-fold) than the endogenous levels, yet no rise in tumors were observed. The lack of an increase in cancer incidence suggests the biological implausibility of a steep initial linear slope as predicted by the USEPA's 2-piece spline model. The ethylene data are conservatively consistent with a shallower slope predicted by the TCEQ's CPH model.

In addition to our comments above, USEPA should consider detailed comments from ACC (2020). ACC (2020) commented to USEPA MON that normal human background endogenous EtO exposure from biological processes equals to about 2 ppb, while smokers' exposure is roughly 20 ppb per day. ACC (2020) argued that if the USEPA's dose-response model were correct, it should show clear cancer signals in smokers, but lymphoid and breast cancers have not been causally linked to smoking.

The USEPA RTC (2022) found that the HEV-adduct method was validated only for ppm-level occupational exposures, and its use at low ppb levels lacked validation. They noted potential for "forward" validation if EtO in smoker breath showed a linear correlation to HEV adducts. ACC (2023) responded with a validation study using published data on EtO in cigarette smoke, demonstrating a continuous linear relationship between low-ppb smoking exposures and high-ppm occupational exposures (Kirman et al., 2025; Figure 4). This validation shows that EtO toxicokinetics remain linear from low to high exposures in humans, supporting PBPK analysis results. It also verifies that smokers do not exhibit cancer signals predicted by the USEPA's two-linear spline model.

Such a conclusion offers an additional reality check that, if the 2-linear spline model were correct, facility-contributed exposures even at a 10^{-4} cancer risk action level (10 ppt) would be completely nested within the noise of normal population variability associated with 2,000 ppt endogenous background exposures. Thus, the USEPA risk model as proposed lacks utility in differentiating true additive cancer risks from those possibly attributed to normal endogenous EtO production. To the current day, USEPA has not commented on this important validation of the smoking low exposure model and its implications as a reality check challenge to the USEPA 2-linear spline model.

5.2. Endogenous Levels as an Important Reality Check for Risk Assessment and Risk Management Considerations

Several questions were posed by ACC (2025) comments on the USEPA Proposed Amendments to NESHAP under the Clean Air Act (CAA) that apply to the proposed CMAS. The following summarizes our response to these questions and why they are important for USEPA to address.

5.2.1 What is the basis for USEPA's concern regarding the reliability of background EtO concentrations in light of new data using USEPA-refined methods? How does USEPA plan to consider total background concentration exposure in managing EtO risk?

USEPA has noted that ambient EtO concentrations near facilities could be probative on the anticipated impacts of EtO emissions on residents (ref from CMAS comments?). However, USEPA declined to account for background EtO concentrations from sources unrelated to industrial emissions due to its belief that there is substantial uncertainty in detection of EtO distant from facilities due to insensitive and unreliable historic test methods. In doing so, USEPA ignores its substantial refinements of canister-based EtO analytical methods for quantifying ambient EtO near and far from EtO facilities, and infers without further analysis, that near-facility concentrations are substantially greater than ambient background concentrations and pose a health issue. This approach fails to acknowledge substantial new data published since the USEPA IRIS indicating everyone is exposed to exogenous background EtO concentrations, as well as importantly to substantially higher endogenous background concentrations (Section 4.1).

State health agencies are collecting EtO data relying on the newly refined USEPA method, and those new data continue to demonstrate a general pattern of many near facility location concentrations being indistinguishable from ambient background concentrations. These data and conclusions have now been published (Kirman et al., 2025). Demonstration of EtO facility-contributed exposures that are clearly (statistically) differentiated from total background exposures is an important initial step in EtO risk management decisions. Particularly important to this consideration is that everyone has an average daily naturally produced endogenous background exposure approximately 4 orders of magnitude greater than the USEPA-modeled 10^{-6} 0.0001 ppb risk specific concentration (Section 4.1). This key risk management consideration is echoed by USEPA itself in its Report to Congress addressing Residual Risk considerations (USEPA, 1999), in which it is explicitly noted "If the relative contribution of background risk is high compared to the incremental residual risk, additional source risk reduction may provide relatively negligible benefit" (at p.94).

USEPA has not provided any rationale for ignoring endogenous EtO levels as a reality check for selection of dose-response models and as a consideration for risk management decisions.

5.2.2. How does USEPA reconcile a potency estimate that suggests that EtO is a highly potent carcinogen at levels substantially below that which the body produces through natural processes? What scientific rationale does USEPA have for ignoring endogenous EtO levels as a reality check for selection of dose-response models and as a consideration for risk management decisions?

USEPA dismisses multiple lines of converging evidence that everyone in the U.S. is exposed to EtO from their own metabolism at concentrations far greater than the USEPA RSC (Kirman et al., 2021; 2025). These new data provide important reality checks in guiding selection of biologically plausible dose-response models appropriate for reliable evaluation of the NIOSH EtO epidemiology dataset. Critically, consideration of substantial endogenous EtO production indicates there is no value in attempting to manage USEPA model-inferred risks from ambient air exposures that are orders of magnitude lower than endogenous equivalent values.

USEPA's primary response to the endogenous EtO reality check is that its cancer risk value estimates extra risk above background and that endogenous levels are intrinsically included in its risk estimates derived from the NIOSH epidemiology cohort. In this regard, USEPA is correct in noting that the contribution of approximate 0.002 ppm daily endogenously generated EtO exposure is indeed trivial and intrinsically nested within risk estimates extrapolated from much higher ppm-level occupational exposures. However, this contention logically collapses when the USEPA estimate of general population risk (10^{-6} risk = 0.0000001 ppm) is far less than even 1% of comparative equivalent daily EtO exposure contributed by endogenous generation.

5.2.3. What is the scientific basis for US EPA's continued skepticism on linear toxicokinetics for EtO over the entire low- to mid-exposure range of NIOSH-cohort occupational exposures, when in fact by its own 2005 Carcinogen Risk Assessment Guidelines low-dose linearity serves as the default assumption for toxicokinetics (i.e., below metabolic saturation, co-factor depletion), and also serves as the default assumption for toxicodynamic factors for low-dose risk from exposure to chemical carcinogens?

The USEPA RTC (2024) maintains that quantitation of EtO exposures based on measurement of EtO-induced hydroxyethyl-hemoglobin (HEV) adduct at ambient exposure levels or of endogenous levels are unreliable. This position is based on USEPA's conclusion that the simple linear model used to develop the HEV versus EtO exposure correlation at low environmental exposures required extrapolation well below occupational exposure levels used in the exposure model development. However, ACC (2023) provided new evidence to USEPA (published as Kirman et al., 2025) based on NHANES biomonitoring data in smokers and in non-smokers that USEPA RTC (2024) did not address. There are only three simple assumptions made by Kirman and Hayes (2017) and Kirman et al. (2021) in development of the EtO HEV toxicokinetic model: (1) The high quality CDC NIOSH NHANES HEV data can be used to quantify *total* (exogenous and endogenous) exposures to EtO; (2) USEPA's air monitoring data can be used to quantify exogenous exposure to EtO in air; and (3) a continuously linear relationship between EtO exposures and HEV measurements is demonstrated between ppm-level occupational exposures down to low-ppb EtO exposures contributed by smoking. The validity of the three considerations has been affirmed by the refined smoking validation analyses published in Kirman et al. (2025). It is important to note that the Kirman et al. (2025) validation approach closely paralleled the USEPA (2022, at p.69) recommendation that such a continuously linear relationship between HEV and EtO could be "forwards" validated by comparative inclusion of

documented low EtO concentrations in cigarette smoke and associated corresponding HEV levels. Despite being presented with this validation effort in 2023 (ACC, 2023), the most recent USEPA RTC (2024) continues to dismiss a continuously linear EtO toxicokinetic dose-response model as hypothetical, and erroneously asserts that no new data on EtO in cigarette smoke had been presented.

Additional evidence supportive of EtO continuously linear toxicokinetics over a wide range of low-to-high exposures has been recently published in mice (Liu et al. 2026). These data show continuously linear toxicokinetics of EtO as measured by HEV over an extremely wide exposure range (0.05 - 50 ppm). EtO blood concentrations transitioned to nonlinear toxicokinetic behavior beginning between 50 and 100 ppm, i.e., exhibited a dose-disproportionate upward bend in HEV between these exposures and continuing up to 200 ppm. These same mice also exhibited linear performance in formation of non-mutagenic N7 DNA adducts (regarded as biomarker of EtO dose delivery to DNA) up to 50 ppm, and again exhibited the same upward bend in adducts beginning between at or after 50 ppm up to the 200 ppm top exposure. EtO genotoxicity and mutagenicity were also apparent only between 100 and 200 ppm, indicating EtO is at best a weakly active genotoxicant that conservatively operates by single linear and shallow dose-response from very low to very high EtO exposures. This direct animal evidence of high-dose specific transition to nonlinear toxicokinetics and mode of action informative DNA adducts and apical genotoxicity is consistent with onset of substantial PBPK-modeled depletion of liver glutathione, the predominant cellular defense contributing to detoxification of EtO in mice, that likewise begins at approximately 100 ppm EtO (Fennell and Brown, 2001). The PBPK model also illustrated that mouse, rat and critically human EtO blood concentrations exactly paralleled and overlaid each other at all concentrations up to 100 ppm, and offered direct countering evidence to the hypothesized supralinear dose-response model that USEPA extrapolates from the NIOSH epidemiology data.

5.2.4. What concrete evidence does USEPA have to support USEPA's current vague claim of "uncertainty" based on some as-yet-to-be-identified source of nonlinear toxicokinetics? Does USEPA agree that the analysis provided by ACC (2023) is new analysis that addresses USEPA (2022) recommendation?

As noted in the answer to the preceding question, ACC's comments on the HON (ACC, 2023; and now Kirman et al., 2025) provide direct experimental validation of the Kirman et al. (2021) linear correlation of EtO blood concentrations from very low to very high level EtO exposures as reflected by HEV adduct data. This correlation was extended to provide further evidence supporting the reliability of estimation of endogenous EtO-equivalent levels (Kirman et al., 2025). In light of this strong new evidence available to USEPA, as well as recent mode of action investigations in mice supporting the biological plausibility only of a single-linear EtO dose response (Liu et al. 2026), the burden of proof is on USEPA to reject the now well supported

and biologically plausible evidence for use of a continuously linear EtO dose-response model to estimate EtO cancer risk.

As described in detail above in Section 5.2.2., EOSA contends that the USEPA is improperly dismissing multiple lines of converging evidence demonstrating that every individual in the U.S. is exposed to EtO via their own metabolism at concentrations significantly exceeding the USEPA's Risk Specific Concentration (RSC). These data (Kirman et al., 2021; 2025) provide a necessary biological reality check that must guide the selection of dose-response models if we are to achieve a reliable evaluation of the NIOSH EtO epidemiology dataset.

There is no regulatory value in attempting to manage model-inferred risks from ambient air exposures that are orders of magnitude lower than these endogenous equivalent value.

5.3. In light of the USEPA IRIS (D29-D30; D40-42) use of statistical methods to select the knot for breast cancer incidence and lymphoid mortality, what is the basis for USEPA's claim that it is not a parameter? What reference does USEPA rely upon to support its approach?

ACC (2020, 2024) and TCEQ (2020) provided detailed comments and references regarding why USEPA's model p-values are incorrect. USEPA claims that the USEPA SAB (2016) supports the USEPA IRIS method. However, this reference does not support what USEPA actually did in their analysis to select the knot.

USEPA's response to comments state the following:

> *"Arguments that the USEPA's fit statistics should be "corrected" rest on an incorrect inference that the USEPA utilized a statistically optimized (maximum likelihood) estimate of the knot value. However, the IRIS assessment clearly indicates that the 1600 ppm-day knot value was not the maximum likelihood estimate of the knot for the lymphoid cancer model - instead, the maximum likelihood knot was at 100 ppm-days (IRIS Appendix D, pg. D-60). If the USEPA had used the maximum likelihood knot estimate of 100 ppm-days, substantially higher estimates of risk would have been obtained. Instead, the USEPA exercised judgment in selecting the model with knot at 1600 ppm-days. The USEPA judged the 1600 ppm-day model as providing a better visual fit to the lower dose data on lymphoid cancer than the maximum likelihood knot of 100 ppm-days. This distinction is important. The fit statistic calculations recommended by commenters assume that the selected parameter estimates are maximum likelihood optimal. As that was not the case in the USEPA modeling, it is incorrect to assert that the USEPA fit statistics should be adjusted using methods that assume that the USEPA's knot value was a mathematical optimum. Also, the USEPA did not use a maximum likelihood knot value in the IRIS assessment*
>
> *"(p.63) Regarding the above SAB advice, the USEPA notes that for lymphoid cancer, the USEPA did fix (i.e., hold constant) the knot value at 1600 ppm-days and, of primary*

importance, determined the upper bound estimate of risk while holding the knot at 1600 ppm-days. If the USEPA had instead allowed the knot to vary in this calculation, the upper bound risk estimate would have been substantially higher - as evidenced by higher risk estimates using the maximum likelihood knot estimate of 100 ppm-days as discussed above. The USEPA also notes that it's draft risk assessment, which was reviewed by the SAB and many public commenters, was clear about its approach of examining a range of knot values before making a selection [US EPA, August 2014. Evaluation of the Inhalation Carcinogenicity of Ethylene Oxide, Revised External Peer Review Draft (USEPA/635/R-14/194b). Appendix D, Table D3-g, pg. D-49.]. The USEPA also notes that the SAB made two drafts of its advice on model selection available for public comment and these drafts both contained recommendations to fixing value of the knot [Science Advisory Board "Draft Report (January 7, 2015) to Assist Meeting Deliberations", pg. 12 and "Draft Report (April 27, 2015) for SAB Quality Review", pg. 13).] There was no lack of clarity about the Board's advice to the USEPA in these public comment drafts.

"Contrary to the commenter's statement, the USEPA notes that fixing the parameter value means to holding it constant through subsequent analyses. For lymphoid cancer the USEPA did fix the knot value at 1600 ppm-days and, of primary importance, determined the upper bound estimate of risk while holding the knot at 1600 ppm-days. If the USEPA had instead allowed the knot to vary in this calculation, the upper bound risk estimate would have been substantially higher - as evinced by higher risk estimates using the maximum likelihood knot estimate of 100 ppm-days as discussed above."

The USEPA's response is misleading because it fails to disclose that the 1,600 ppm-day knot in its preferred model was identified as a "local maximum likelihood" knot via a systematic 100-unit grid-search optimization for lymphoid mortality cancers. Despite rejecting the global maximum likelihood knot of 100 ppm-days as "biologically implausible," the USEPA used this optimization to find its second-choice knot based on the local maximum likelihood, while publicly attributing the selection to "visual fit." Because the knot was an estimated parameter, standard statistical practice requires the subtraction of one degree of freedom—a step the USEPA bypassed by claiming the selection occurred in a preliminary stage.

The USEPA's refusal to reduce degrees of freedom is arbitrary, statistically unsound, and based on a mischaracterization of SAB recommendations. While the SAB noted that fixing parameters can be appropriate, it did not authorize violating the fundamental rule that parameters estimated from data must be accounted for in the model's degrees of freedom and AIC calculations.

The USEPA IRIS (2016) further demonstrates a lack of consistency: for breast cancer incidence and mortality, the agency selected the maximum likelihood knot (Appendix D-8, D-18) and presented p-values both with and without accounting for it as a parameter (Appendix D-13). This inconsistency suggests the USEPA either misunderstood its own analysis or is misrepresenting it.

By ignoring standard statistical practices, the USEPA's approach lacks scientific credibility and relies on a revisionist interpretation of SAB advice.

5.4. Does USEPA agree that USEPA (2024) Table 1 compares the rate ratios (relative rate) that reflect those on Figure 4-3, and that this analysis is the same as comparing the models along the y-axis? If not, why not? How does USEPA explain this approach in light of USEPA (IRIS) Figure 4-3 footnote warning against such comparisons? What references or citations does USEPA have to refute or fail to adhere to the USEPA (IRIS) 2016 Figure 4-3 footnote warning?

The USEPA's (2024) RTC Table 1 introduces a technically contradictory critique of the TCEQ's model fit by claiming it underestimates NIOSH categorical risk estimates. By approximating internal rate ratio (RR) values for each average exposure quartiles from the continuous models, the USEPA performs a y-axis comparison that is explicitly prohibited by a cautionary note in its own IRIS Figure 4-3.

USEPA's analysis in Table 1 relies on a fundamental misapplication of visual diagnostics, specifically by performing vertical (y-axis) comparisons between models that are independently normalized. All continuous models are anchored such that the Relative Risk (RR) is exactly 1.0 at zero cumulative exposure. Because each model calculates its slope and curvature based on different equations fit through individual data (e.g., linear, log-linear), the y-axis values are relative only to each model's own internal baseline.

The statistical implication of the USEPA's approach is a **scaling fallacy**: comparing the absolute vertical distance between a model curve and a categorical data point is mathematically invalid when the curves are not on a shared absolute scale. Since the y-intercept is a fixed point (RR=1) by definition, the vertical "spread" between models at higher exposures is a reflection of their different mathematical shapes, not an indicator of "error" or "poor fit." By ignoring the note in Figure 4-3 and focusing on these relative y-axis offsets, the USEPA fails to acknowledge that model fit must be evaluated based on the likelihood of the entire individual dataset, rather than the visual proximity to grouped, categorical proxies along the y-axis.

5.5. Does USEPA agree that applying the CPH model to each categorical estimate using one categorical modeled estimate at a time is incorrect and irrelevant to TCEQ's analysis? If not, why not and what literature supports USEPA's approach? Does USEPA agree that applying the same approach to the linear spline model focusing on the initial linear CPH model below the knot would result in similar overprediction of the categorical estimates at later exposures? Does USEPA agree that this approach is diametrically opposite from SAB's strong recommendation to NOT model the categorical data? If not, why not?

USEPA's (2022, 2024) Response to Comments (RTC) employs a misleading visual assessment of lymphoid categorical data to claim the TCEQ's log-linear Cox Proportional Hazards (CPH)

model "explodes" at high cumulative exposures. This critique is technically flawed, as applying a CPH model to grouped estimates one at a time is mathematically irrelevant to TCEQ's actual analysis, which models all the individual-level data. Furthermore, a direct comparison of rate ratios at high-exposure benchmarks (40,000 and 64,000 ppm-days) reveals that the USEPA's model actually yields higher predicted risks than the TCEQ's. The USEPA mischaracterizes the TCEQ's approach as uniquely extreme, while its own model maintains a steeper risk profile across the same range. Finally, the USEPA's reliance on categorical data for model evaluation is diametrically opposed to the Science Advisory Board's (SAB) explicit recommendation to prioritize continuous, individual data over grouped categorical estimates.

References

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

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

Berman, N. G., Wong, W. K., Bhasin, S., Ipp, E. (1996). Applications of segmented regression models for biomedical studies. *The American journal of physiology*, *270*(4 Pt 1), E723–E732. https://doi.org/10.1152/ajpendo.1996.270.4.E723Bogen, K. T., Sheehan, P. J., Valdez-Flores, C., & Li, A. 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. *International journal of environmental research and public health*, *16*(10), 1738. https://doi.org/10.3390/ijerph16101738

Bruch CW. (1972). Sterilization of plastics: Toxicity of ethylene oxide residues. In Industrial Sterilization, Proceedings of the International Symposium, Amsterdam, The Netherlands. Phillips, G.B., Miller, W.S., Eds.; *Duke University Press,* Durham, NC, USA.

Centers for Disease Control and Prevention. (n.d.). *National report on human exposure to environmental chemicals: Ethylene oxide hemoglobin adducts*. Available online: https://www.cdc.gov/exposurereport/index.html

Fennell, T. R., Brown, C. D. (2001). A physiologically based pharmacokinetic model for ethylene oxide in mouse, rat, and human. *Toxicology and applied pharmacology*, *173*(3), 161–175. https://doi.org/10.1006/taap.2001.9184

Filser, J. G., & Klein, D. (2018). A physiologically based toxicokinetic model for inhaled ethylene and ethylene oxide in mouse, rat, and human. *Toxicology letters*, *286*, 54–79. https://doi.org/10.1016/j.toxlet.2017.07.896

Goldgraben R, Zank N. (1981). Mitigation of Worker Exposure to Ethylene Oxide. Report prepared for the US Environmental Protection Agency, MTR-80 W333 (Sponsor: U.S. Environmental Protection Agency, Contract No.: 68-01·5944). The Mitre Corp: McLean, VA. https://nepis.epa.gov/Exe/ZyPURL.cgi%3FDockey%3D940000ON.TXT&ved=2ahUKEwiDksG2iMSTAxU9KVkFHRklMHYQFnoECBgQAQ&usg=AOvVaw1COZjiTNiuQVn3QqR1r6om

Gollapudi, B. B., Bus, J. E., Cassidy, P., Weinberg, J. T., Bemis, J. C., Torus, D. T., Dertinger, S. D., Lu, K., & Li, A. P. (2026). Investigation of ethylene oxide genotoxicity dose-response to inform cancer risk assessment. *BioRxiv*.

Gollapudi, B. B., Su, S., Li, A. A., Johnson, G. E., Reiss, R., & Albertini, R. J. (2020). Genotoxicity as a toxicologically relevant endpoint to inform risk assessment: A case study with ethylene oxide. *Environmental and molecular mutagenesis*, *61*(9), 852–871. https://doi.org/10.1002/em.22408

Greife, A. L., Hornung, R. W., Stayner, L. G., & Steenland, K. N. (1988). Development of a model for use in estimating exposure to ethylene oxide in a retrospective cohort mortality study. *Scandinavian journal of work, environment & health*, *14 Suppl 1*, 29–30.

Hamm, T. E., Jr, Guest, D., Dent, J. G. (1984). Chronic toxicity and oncogenicity bioassay of inhaled ethylene in Fischer-344 rats. *Fundamental and applied toxicology : official journal of the Society of Toxicology*, *4*(3 Pt 1), 473–478. https://doi.org/10.1016/0272-0590(84)90205-7

Heflich, R. H., Johnson, G. E., Zeller, A., Marchetti, F., Douglas, G. R., Witt, K. L., Gollapudi, B. B., & White, P. A. (2020). Mutation as a Toxicological Endpoint for Regulatory Decision-Making. *Environmental and molecular mutagenesis*, *61*(1), 34–41. https://doi.org/10.1002/em.22338

Höge M, Wöhling T, Nowak W. A. (2018). primer for model selection: the decisive role of model complexity. *Water Resources Res, 54*(3):1688-1715. doi: 10.1002/2017wr021902.

Hornung, R. W., Greife, A. L., Stayner, L. T., Steenland, N. K., Herrick, R. F., Elliott, L. J., Ringenburg, V. L., & Morawetz, J. (1994). Statistical model for prediction of retrospective exposure to ethylene oxide in an occupational mortality study. *American journal of industrial medicine*, *25*(6), 825–836. https://doi.org/10.1002/ajim.4700250607

International Agency for Research on Cancer (IARC). (2008). *IARC monographs on the evaluation of carcinogenic risks to humans: Volume 97. 1,3-butadiene, ethylene oxide and vinyl halides (vinyl fluoride, vinyl chloride and vinyl bromide)* (pp. 3–471). IARC.

International Commission for Protection Against Environmental Mutagens and Carcinogens (ICPEMC). (1980). ICPEMC news no. 2. Environmental Health Perspectives, 34, 223–225.

Kelly-Reif, K., Bertke, S. J., Stayner, L., Steenland, K. (2025). Exposure to Ethylene Oxide and Relative Rates of Female Breast Cancer Mortality: 62 Years of Follow-Up in a Large US Occupational Cohort. *Environmental health perspectives*, *133*(5), 57013. https://doi.org/10.1289/EHP15566

Kirman, C. R., Bus, J. S. (2026). Regarding ethylene exposure and endogenous ethylene oxide levels in humans (Lin et al., 2025). Submitted to *Journal of Exposure Science and Environmental Epidemiology*.

Kirman, C. R., & Hays, S. M. (2017). Derivation of endogenous equivalent values to support risk assessment and risk management decisions for an endogenous carcinogen: Ethylene

oxide. *Regulatory toxicology and pharmacology : RTP*, *91*, 165–172. https://doi.org/10.1016/j.yrtph.2017.10.032

Kirman, C. R., Li, A. A., Sheehan, P. J., Bus, J. S., Lewis, R. C., & Hays, S. M. (2021). Ethylene oxide review: characterization of total exposure via endogenous and exogenous pathways and their implications to risk assessment and risk management. *Journal of toxicology and environmental health. Part B, Critical reviews*, *24*(1), 1–29. https://doi.org/10.1080/10937404.2020.1852988

Kirman, C. R., Sheehan, P. J., Li, A. A., Bus, J. S., Su, S. H., Dopart, P. J., Watson, H. N., Moynihan, E. E., & Reiss, R. (2025). Characterization of Background Exposures to Ethylene Oxide in the United States: A Reality Check on Theoretical Health Risks for Potentially Exposed Populations near Industrial Sources. *International journal of environmental research and public health*, *22*(4), 597. https://doi.org/10.3390/ijerph22040597 7.

Klapacz, J., Pottenger, L. H., Engelward, B. P., Heinen, C. D., Johnson, G. E., Clewell, R. A., Carmichael, P. L., Adeleye, Y., & Andersen, M. E. (2016). Contributions of DNA repair and damage response pathways to the non-linear genotoxic responses of alkylating agents. *Mutation research. Reviews in mutation research*, *767*, 77–91. https://doi.org/10.1016/j.mrrev.2015.11.001

LeBaron, M.J, Schisler, M., Zhang, F., Jeong, Y. (2012). Bioanalytical, Molecular, and Genetic Alterations in Ethylene Oxide-Exposed Male B6C3F1 Mice. S65-S65. *Environmental and Molecular Mutagenesis, 53*, S65.

Lewis, R. C., Sheehan, P. J., DesAutels, C. G., Watson, H. N., & Kirman, C. R. (2022). Monitored and Modeled Ambient Air Concentrations of Ethylene Oxide: Contextualizing Health Risk for Potentially Exposed Populations in Georgia. *International journal of environmental research and public health*, *19*(6), 3364. https://doi.org/10.3390/ijerph19063364

Li, W., He, C., & Freudenberg, J. (2011). A mathematical framework for examining whether a minimum number of chiasmata is required per metacentric chromosome or chromosome arm in human. *Genomics*, *97*(3), 186–192. https://doi.org/10.1016/j.ygeno.2010.11.007

Lin, Y. S., Morozov, V., & Wu, K. Y. (2026). Integrating PBPK modeling and tobacco biomarkers to interpret N-(2-hydroxyethyl)valine (HEV) hemoglobin adducts in the U.S. population. *International journal of hygiene and environmental health*, *272*, 114721. https://doi.org/10.1016/j.ijheh.2025.114721

Lin, Y. S., Thayer, K. A., White, P., Morozov, V., & Persad, A. S. (2025). Uncovering the connection: ethylene exposure and endogenous ethylene oxide levels in humans. *Journal of exposure science & environmental epidemiology*, *36*(2), 361–374. https://doi.org/10.1038/s41370-025-00826-7

Liu, C.-W., Peng, J., Feng, J., Zhao, H., Wang, X., Gollapudi, B., Li, A., Bus, J., & Lu, K. (2026). Molecular dosimetry of DNA adducts in mice exposed to ethylene oxide. *BioRxiv*. https://www.biorxiv.org/content/10.64898/2026.03.25.714191v1.

Lynch, D. W., Lewis, T. R., Moorman, W. J., Burg, J. R., Groth, D. H., Khan, A., Ackerman, L. J., & Cockrell, B. Y. (1984). Carcinogenic and toxicologic effects of inhaled ethylene oxide and propylene oxide in F344 rats. *Toxicology and applied pharmacology*, *76*(1), 69–84. https://doi.org/10.1016/0041-008x(84)90030-9

MacRae, S. L., Croken, M. M., Calder, R. B., Aliper, A., Milholland, B., White, R. R., Zhavoronkov, A., Gladyshev, V. N., Seluanov, A., Gorbunova, V., Zhang, Z. D., & Vijg, J. (2015). DNA repair in species with extreme lifespan differences. *Aging*, *7*(12), 1171–1184. https://doi.org/10.18632/aging.100866

Manjanatha, M. G., Shelton, S. D., Chen, Y., Parsons, B. L., Myers, M. B., McKim, K. L., Gollapudi, B. B., Moore, N. P., Haber, L. T., Allen, B., & Moore, M. M. (2017). Dose and temporal evaluation of ethylene oxide-induced mutagenicity in the lungs of male big blue mice following inhalation exposure to carcinogenic concentrations. *Environmental and molecular mutagenesis*, *58*(3), 122–134. https://doi.org/10.1002/em.22080

Marsh, G. M., Keeton, K. A., Riordan, A. S., Best, E. A., & Benson, S. M. (2019). Ethylene oxide and risk of lympho-hematopoietic cancer and breast cancer: a systematic literature review and meta-analysis. *International archives of occupational and environmental health*, *92*(7), 919–939. https://doi.org/10.1007/s00420-019-01438-z

Molinari, N., Daurès, J. P., & Durand, J. F. (2001). Regression splines for threshold selection in survival data analysis. *Statistics in medicine*, *20*(2), 237–247. https://doi.org/10.1002/1097-0258(20010130)20:2<237::aid-sim654>3.0.co;2-i

Moore, M. M., Heflich, R. H., Haber, L. T., Allen, B. C., Shipp, A. M., & Kodell, R. L. (2008). Analysis of in vivo mutation data can inform cancer risk assessment. *Regulatory toxicology and pharmacology : RTP*, *51*(2), 151–161. https://doi.org/10.1016/j.yrtph.2008.01.015

National Academies of Sciences, Engineering, and Medicine. (2025). *Review of Texas Commission on Environmental Quality's Ethylene Oxide Development Support Document*. Washington, DC: National Academies Press. ISBN-13: 978-0-309-73165-2; ISBN-10: 0-309-73165-8.

National Research Council (NRC). (2007). Models in Environmental Regulatory Decision Making. Chapter 5: Model Selection and Use (pp. 170-189). National Academies Press, Washington, DC.

National Toxicology Program (1987). NTP Toxicology and Carcinogenesis Studies of Ethylene Oxide (CAS No. 75-21-8) in B6C3F1 Mice (Inhalation Studies). *National Toxicology Program technical report series*, *326*, 1–114.

Parsons, B. L., Beal, M. A., Dearfield, K. L., Douglas, G. R., Gi, M., Gollapudi, B. B., Heflich, R. H., Horibata, K., Kenyon, M., Long, A. S., Lovell, D. P., Lynch, A. M., Myers, M. B., Pfuhler, S., Vespa, A., Zeller, A., Johnson, G. E., & White, P. A. (2025). Severity of effect considerations regarding the use of mutation as a toxicological endpoint for risk assessment: A report from the 8th International Workshop on Genotoxicity Testing (IWGT). *Environmental and molecular mutagenesis*, *66 Suppl 2*(Suppl 2), 121–143. https://doi.org/10.1002/em.22599

Picciotto S, Kelly-Reif K, Eisen EA, Stayner LT, Costello S. How to identify the healthy worker survivor effect empirically and how to interpret results from published studies: the NIOSH ethylene oxide cohort as a case study. Am J Epidemiol. 2026 Mar 9:kwag052. doi: 10.1093/aje/kwag052. Epub ahead of print. PMID: 41800785.

Recio, L., Donner, M., Abernethy, D., Pluta, L., Steen, A. M., Wong, B. A., James, A., & Preston, R. J. (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*(3), 215–222. https://doi.org/10.1093/mutage/geh017

Rodríguez-Domínguez R, Castillo-Vargasmachuca SG, Pérez-González R, Aragón-Noriega EA. (2018). Allometry in *Callinectes bellicosus* (Stimpson, 1859) (Decapoda: Brachyura: Portunidae): single-power model versus multi-model approach. *J Crutacean Biol, 38*(5):574-78. doi: 10.1093/jcbiol/ruy060.

Roy, R., Shiota, S., Kennel, S. J., Raha, R., von Wronski, M., Brent, T. P., & Mitra, S. (1995). A comparative study of the biochemical properties of human and mouse recombinant O6-methylguanine-DNA methyltransferases. *Carcinogenesis*, *16*(2), 405–411. https://doi.org/10.1093/carcin/16.2.405

Sheehan, P. J., Lewis, R. C., Kirman, C. R., Watson, H. N., Winegar, E. D., & Bus, J. S. (2021). Ethylene Oxide Exposure in U.S. Populations Residing Near Sterilization and Other Industrial Facilities: Context Based on Endogenous and Total Equivalent Concentration Exposures. *International journal of environmental research and public health*, *18*(2), 607. https://doi.org/10.3390/ijerph18020607

Sheehan, P. J., Lewis, R. C., Kirman, C. R., Watson, H. N., Winegar, E. D., & Bus, J. S. (2021). Ethylene Oxide Exposure in U.S. Populations Residing Near Sterilization and Other Industrial Facilities: Context Based on Endogenous and Total Equivalent Concentration Exposures. *International journal of environmental research and public health*, *18*(2), 607. https://doi.org/10.3390/ijerph18020607

Sielken RL, Valdez Flores C. (2009a). Life-table calculations of excess risk for incidence versus mortality: ethylene oxide case study. *Regul Toxicol Pharmacol 55*(1):82-89.

Snellings, W. M., Weil, C. S., & Maronpot, R. R. (1984). A two-year inhalation study of the carcinogenic potential of ethylene oxide in Fischer 344 rats. *Toxicology and applied pharmacology*, *75*(1), 105–117. https://doi.org/10.1016/0041-008x(84)90081-4

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. *Occupational and environmental medicine*, *61*(1), 2–7.

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. https://doi.org/10.1056/NEJM199105163242004

Steenland, K., Whelan, E., Deddens, J., Stayner, L., & Ward, E. (2003). Ethylene oxide and breast cancer incidence in a cohort study of 7576 women (United States). *Cancer causes & control : CCC*, *14*(6), 531–539. https://doi.org/10.1023/a:1024891529592

Stetson, J. B., Whitbourne, J. E., & Eastman, C. (1976). Ethylene oxide degassing of rubber and plastic materials. *Anesthesiology*, *44*(2), 174–180. https://doi.org/10.1097/00000542-197602000-00023

Texas Commission on Environmental Quality (TCEQ). (2020). *Ethylene oxide carcinogenic dose-response assessment; CAS Registry Number: 75-21-8; In Development Support Document Final, May 15, 2020*. Texas Commission on Environmental Quality: Austin, TX, USA.

Troy, J. D., Hartge, P., Weissfeld, J. L., Oken, M. M., Colditz, G. A., Mechanic, L. E., & Morton, L. M. (2010). Associations between anthropometry, cigarette smoking, alcohol consumption, and non-Hodgkin lymphoma in the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial. *American journal of epidemiology*, *171*(12), 1270–1281. https://doi.org/10.1093/aje/kwq085

U.S. Environmental Protection Agency (USEPA SAB). (2015). *Science Advisory Board Review of the USEPA's Evaluation of the Inhalation Carcinogenicity of Ethylene Oxide (Revised External Review Draft—August 2014); USEPA-SAB-15-012 (2015 SAB Review)*. Washington, DC.

U.S. Environmental Protection Agency (USEPA). 2005. *Guidelines for Carcinogen Risk Assessment.* USEPA/630/P-03/001F. US EPA Risk Assessment Forum, Washington, DC.

U.S. Environmental Protection Agency (US EPA). (2022). *Summary of public comments and responses for the reconsideration of the 2020 National Emission Standards for Hazardous Air Pollutants: Miscellaneous Organic Chemical Manufacturing Residual Risk and Technology Review*. Office of Air Quality Planning and Standards Sector Policies and Programs Division (EE-143-01), Research Triangle Park, NC 27711.

U.S. Environmental Protection Agency (US EPA). (2024). *Summary of Public Comments and Responses for New Source Performance Standards for the Synthetic Organic Chemical Manufacturing Industry and National Emission Standards for Hazardous Air Pollutants for the Synthetic Organic Chemical Manufacturing Industry and Group I & II Polymers and Resins Industry.* Office of Air Quality Planning and Standards, Research Triangle Park, NC. Docket ID

No. USEPA-HQ-OAR-2024-0303, posted by US EPA May 16, 2024. https://www.regulations.gov/document/USEPA-HQ-OAR-2022-0730-2764

U.S. Environmental Protection Agency, Integrated Risk Information System (US EPA IRIS). (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)*. U.S. Environmental Protection Agency, Office of Research and Development, National Center for Environmental Assessment: Washington, DC.

U.S. Environmental Protection Agency. (2008). *U.S. Environmental Protection Agency. Reregistration Eligibility Decision for Ethylene Oxide. USEPA 738-R-08-003.* U.S. Environmental Protection Agency, Office of Pesticide Programs. Washington, DC. March.

U.S. Environmental Protection Agency (2026). Future development and use of risk assessment. Memo from David Fotuthi, EPA General Counsel.

U.S. Food and Drug Administration (USFDA). (1978). *Ethylene Oxide, Ethlyene Chlorohydrin, and Ethylene Glycol: Proposed Maximum Residue Limits and Maximum Levels of Exposure. Federal Regist 43(122):27473-83*; US Food and Drug Administration: Washington, DC, USA.

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

Valdez-Flores, C., Li, A. A., Bender, T. J., & Teta, M. J. (2025). Use of updated mortality study of ethylene oxide manufacturing workers to inform cancer risk assessment. *Risk analysis : an official publication of the Society for Risk Analysis*, *45*(9), 2822–2837. https://doi.org/10.1111/risa.70057

Valdez-Flores, C., & Sielken, R. L., Jr (2013). Misinterpretation of categorical rate ratios and inappropriate exposure-response model fitting can lead to biased estimates of risk: ethylene oxide case study. *Regulatory toxicology and pharmacology: RTP*, *67*(2), 206–214. https://doi.org/10.1016/j.yrtph.2013.07.011

Walker, V. E., Wu, K. Y., Upton, P. B., Ranasinghe, A., Scheller, N., Cho, M. H., Vergnes, J. S., Skopek, T. R., & Swenberg, J. A. (2000). Biomarkers of exposure and effect as indicators of potential carcinogenic risk arising from in vivo metabolism of ethylene to ethylene oxide. *Carcinogenesis*, *21*(9), 1661–1669. https://doi.org/10.1093/carcin/21.9.1661

White J. D. (1977). Standard aeration for gas-sterilized plastics. *The Journal of hygiene*, *79*(2), 225–232. https://doi.org/10.1017/s0022172400053031

White, P. A., Long, A. S., & Johnson, G. E. (2020). Quantitative Interpretation of Genetic Toxicity Dose-Response Data for Risk Assessment and Regulatory Decision-Making: Current

Status and Emerging Priorities. *Environmental and molecular mutagenesis*, *61*(1), 66–83. https://doi.org/10.1002/em.22351

Volume 3

TABLE OF CONTENTS

Technical Comments

Question #7: Should the EPA modify the emission reductions levels for new ARVs at facilities where EtO use is at least 10 tpy from 99.9 to 99.6 percent reduction, as proposed?

Question #8: Should the EPA consider manufacturer guarantee levels of EtO reduction for ARVs, and if so, what are these levels?

In response to the EPA's questions 7 and 8, we provide the specific comments below.

While we support the EPA's pragmatic proposal to reduce the required emission reduction level for new Aeration Room Vents (ARVs) at facilities utilizing at least 10 tons per year (tpy) of EtO, a strict 99.6% reduction mandate remains technically prohibitive without corresponding regulatory flexibility, *e.g.,* an alternative outlet concentration. Moving away from the unachievable 99.9% requirement is a necessary step; however, enforcing a strict 99.6% reduction standard without an alternative option (*e.g.,* an alternative outlet concentration limit) still exceeds what is commercially demonstrable in many operational scenarios.

Emission standards must be based on established technology and achievable, Original Equipment Manufacturer (OEM) guaranteed destruction and removal efficiency of emission control equipment, with an alternative outlet concentration, to ensure compliance can be achieved and demonstrated. An appropriate, measurable alternative concentration standard for ARVs is imperative. Pre-treatment ARV concentrations can vary significantly and concentrations as low as single-digit ppm or lower are common.

Actual technical constraints making it extremely difficult, if not impossible, for facilities to meet the proposed ARV Destruction and Removal Efficiencies (DREs), without an alternative option:

Discrepancy with Commercially Available Manufacturer Guarantees: Although 99.6% represents a less stringent requirement than the previous 99.9% threshold, it continues to conflict with standard, commercially available equipment guarantees for different emission control technologies. Leading emission control technology manufacturers (*e.g.,* those producing catalytic oxidizers, gas scrubbers, dry scrubbers, and thermal oxidizers) typically only guarantee a 99.0% DRE when the inlet EtO is above a certain threshold (*e.g.,* 1 ppm). As a general rule, the higher the inlet EtO concentration the higher the DRE that can be achieved, and vice versa. While technologies like catalytic oxidation can achieve a 99.9%+ DRE under peak loads, their efficiency is significantly curtailed at low inlet concentrations, where outlet concentrations often fall below the detection limit. Manufacturers often only guarantee a DRE of 99% when inlet EtO is above a

certain threshold (*e.g.,* 1 ppm). For example, as noted in the Picarro and LESNI White Paper,[3] "*to achieve the guaranteed 99.99+% DRE with LESNI's EtO Catalytic Abatement System, it is imperative that customers adhere to the specified operational requirements. These include maintaining consistent EtO inlet concentrations at 50% or above the operating capacity within daily and 30-day rolling sum EtO mass loads along with implementing robust monitoring and reporting solutions such as those offered by Picarro. Deviations from these operational parameters may result in reduced efficiency and non-compliance with regulatory standards. Picarro and LESNI can also only guarantee successful compliance associated with their equipment, and cannot certify that other technologies used for abatement as part of a blended SWEL will always meet their DRE requirements*". LESNI EtO Catalytic Abatement Systems are used to control SCV, ARV, and CEV exhausts, therefore the 50% of operating capacity includes the need for SCVs to be operating at 50% capacity. This LESNI guarantee is not valid for systems only controlling ARV exhausts but represents the reliance on minimum inlet concentrations in guaranteeing efficiency. Another EtO emission control technology provider also notes on its website,[4] that its EtO Scrubber can achieve >99% destruction efficiency for high inlet concentrations, and that its Dry Bed Scrubber systems are highly effective, achieving >99% destruction efficiency for inlet concentrations above ~5 ppm. These guarantees are often contingent upon specific, higher inlet concentrations that do not reflect all operational realities.

Mathematical Impossibility: In addition, enforcing a strict 99.6% percentage-based reduction is exceptionally difficult, if not mathematically impossible, without a corresponding alternative outlet concentration limit. As the inlet concentration drops, calculating a continuous 99.6% DRE becomes practically prohibitive due to instrument detection limits and inherent measurement uncertainty. Furthermore, the EPA's current definition of "Sterilization Operation" includes post-aeration operations, which inherently produce very low inlet concentrations. For example, an inlet concentration of 1 ppm requires the detection system to reliably measure a 4 ppb EtO concentration at the outlet to achieve a 99.6% DRE, and this 4-ppb level is at or below the "representative Detection Level (DL)" of most, if not all, detection systems. As such, providing an alternative outlet concentration limit is essential (see additional information in our comment # 22). Without it, facilities will be unfairly penalized for operating highly efficient control devices simply because the inlet loading is too low to prove a 99.6% reduction mathematically.

CEMS Compliance demonstration at 99.6%: We have significant concerns regarding the ability of Continuous Emission Monitoring Systems (CEMS) to consistently and accurately

---

[3] Picarro and LESNO White Paper. Guaranteeing Subpart-O EtO SWEL1 and SWEL2 Compliance with LESNI EO CAPs and Picarro CEMS. Available at https://www.lesni.com/media/0mdbfupy/whitepaper-20-guaranteeing-suppart-o-long-version.pdf

[4] See https://zjbocon.com/eo-scrubber/

verify a 99.6% DRE for ARVs. The EPA typically mandates CEMS Data Availability of at least 90% to 95%. However, hardware uptime guarantees from manufacturers do not equate to valid data availability at ultra-trace levels. At these near-zero limits, baseline noise and calibration drift significantly expand the uncertainty range of the measurements, severely compromising fundamental accuracy. While this drift may not automatically invalidate the data outright, the resulting margin of error at such ultra-low concentrations risks rendering the data effectively unusable for compliance verification, thereby threatening a facility's ability to meet the EPA's mandatory data availability thresholds. This concern is amplified in the "low EtO concentration scenario" frequently encountered in modern sterilization facilities, particularly those employing advanced pre-conditioning or high-volume room air evacuation. For sources where the inlet concentration is already low (*e.g.,* at or below 1 ppm), the various forms of bias which are reflected in sampling performance standards can create significant accuracy errors, making it difficult to prove actual compliance with a 99.6% DRE, and pushing the required outlet concentration well below the reliable detection limits of current CEMS technology.

In conclusion, while we acknowledge the EPA's shift from a 99.9% to a 99.6% removal efficiency threshold, this adjustment remains insufficient without an alternative outlet concentration limit. The 99.6% DRE requirement for ARVs continues to conflict with standard, commercially available manufacturer guarantees for various emission control technologies, which often cannot certify such high-efficiency percentages across all operational scenarios. Furthermore, the EPA must recognize the absolute necessity of an alternative outlet concentration limit (*e.g.,* < 1 ppm) to provide a viable compliance pathway when inlet concentrations are very low. Without an alternative concentration-based standard, facilities are forced into a "mathematical impossibility" where even negligible emissions fail to meet a percentage-based DRE for ARVs.

Question #9: Should facilities have the option of using either parametric monitoring and performance testing or CEMS to demonstrate initial and continuous compliance with the Commercial Sterilization Facilities NESHAP, as proposed?

In response to the EPA's question 9, we provide the specific comments below.

We support the EPA's proposal to allow facilities the option of using either parametric monitoring and performance testing or CEMS to demonstrate initial and annual compliance. We urge the EPA to establish a compliance framework that prioritizes operational reality and flexibility over rigid monitoring mandates. Below are our specific technical comments.

Allow Hybrid Monitoring via a Flexible SWEL Approach: The EPA must explicitly permit a Site-Wide Emission Limit (SWEL) framework that allows for "hybrid monitoring." Specifically, a facility must be permitted to utilize CEMS for primary, high-load sources (*e.g.,* SCVs) while concurrently utilizing unit-specific parametric monitoring for low-concentration or trace-level systems (*e.g.,* dry bed scrubbers processing room air), where CEMS are mathematically and analytically unsuitable.

To implement this hybrid approach, the EPA must include a clear regulatory mechanism for apportionment or distribution approval. There is strong historical precedent for this flexibility. In the November 2, 2001, Subpart O amendments (66 FR 55577),[5] the EPA allowed facilities with combined, split, or manifolded emission streams to apportion their compliance demonstrations using site-specific engineering data and pre-split performance testing.

Provide Transitional Flexibility and Eliminate "Lock-In" Mandates: Furthermore, the EPA must provide a clear, streamlined administrative pathway for facilities to transition between continuous compliance demonstration methods. If a facility initially elects a unified compliance strategy (*e.g.,* an "all-parametric" or "all-CEMS" SWEL approach), it must not be permanently "locked in" to that initial selection.

Real-world operational data gathered during initial compliance phases may reveal that CEMS are overly susceptible to baseline noise, calibration drift, and false exceedances at ultra-trace levels. If empirical, site-specific data subsequently prove that periodic performance testing combined with parametric monitoring provides a more stable, legally defensible, and accurate emissions profile for a facility's specific configuration, the rule must allow a pivot. Facilities must be afforded the transitional flexibility to switch between CEMS and parametric monitoring based on demonstrated operational stability, without being penalized or forced into protracted permit renegotiations or prenotification requirements for abandoning unworkable monitoring technology.

Question #10: Should the EPA finalize the proposed changes to performance testing, parametric monitoring, and reporting requirements as proposed?

In response to the EPA's question 10, we provide the specific comments below.

We strongly recommend that the EPA simplify the Operating Parameter Limit (OPL) mandates. The EPA's proposal to strictly set OPLs based on conditions captured during the initial and annual performance tests will result in significant operational challenges. Demonstrating the absolute maximum and minimum operational parameters during initial and annual performance testing is extremely difficult, if not technically impossible, for most facilities. Replicating these artificial conditions within the narrow window of a performance test requires a level of environmental and process manipulation that does not reflect actual facility operations. Facilities often cannot artificially control or manipulate certain parameters during the narrow window of a performance test to capture the absolute "maximum" or "minimum" operational ranges they might safely experience over the course of a year. Establishing rigid OPLs based only on a snapshot in time fails to account for normal, compliant operational variability.

Furthermore, the proposed framework is overly prescriptive and treats all control devices with the same level of rigidity, which is impractical. Secondary controls should have less prescriptive OPL requirements, reflecting their role as backup or polishing systems rather than the main reduction technology.

[5] 66 Fed. Reg. 55577

The fundamental goal of the NESHAP is emission reduction. Therefore, the most critical factor for environmental compliance is identifying the key parameters while also considering whether these are primary or secondary control systems. In particular, the EPA must distinguish between parameters that are directly adjustable and those that are purely resultant; for instance, facilities have direct operational control over the burner temperature, but cannot independently control the temperature differential across the catalyst bed. As such, we strongly urge the EPA to simplify the compliance framework to focus on the most important OPLs that would impact EtO emission limits, and allow OPL ranges to be determined using varying sources of information such as performance testing, manufacturer design criteria, and engineering review as opposed to prohibiting any OPL that could not be replicated during a performance test scenario.

A workable suggestion is to return to the simplicity offered in the 1994 EtO NESHAP[6] and focus strictly on critical indicators. For all technologies, intermediate parameters can serve as indicators to identify trends for facility operators, but they do not indicate noncompliant operations that need to be specified as strict, inflexible regulatory limits. Compliance should be determined by critical OPLs with set parameter ranges determined using a multitude of sources combined with testing to verify destruction efficiency, paired with a simplified OPL framework.

In addition, we note a discrepancy regarding the data acquisition requirements on page 61 of the revised redline NESHAP document, which states: "*Continuous monitoring system (CMS) means, for the purposes of this rule, the equipment necessary to continuously sample the regulated parameter specified in § 63.364 or § 63.365 of this subpart without interruption, evaluate the detector response at least once every **15 seconds**, and compute and record the average value at least every **60 seconds**, except during allowable periods of calibration and except as defined otherwise by the continuous emission monitoring system (CEMS) performance specifications (PS) in appendix B to part 60 of this chapter*". Similarly, on page 78 of the revised redline document, it states: *"(v) A data acquisition system capable of collecting a value every **15 seconds** and compute and record each 1-hour block average*". We request that the EPA clarify whether the "15-second" interval is a typographical error intended to be "15 minutes". Requiring a value every 15 seconds to compute a 1-hour block average is inconsistent with the General Provisions of 40 CFR Part 63 and places an unnecessary burden on Data Acquisition and Handling Systems (DAHS) without providing additional environmental benefit. Realigning this text with the standard 15-minute interval would ensure consistency across the NESHAP subparts and prevent confusion during compliance demonstrations and electronic reporting.

We also urge the EPA to reconsider the volume of data mandated for the initial and periodic reporting specifically for CEMS. The current proposal effectively shifts the compliance burden from "maintaining records" to "reporting massive datasets", which complicates the regulatory review process for both the facility and the Agency. For example, the requirement to report every 1-hour block average for multiple parameters related to PTE records creates an excessive administrative burden. We recommend that the EPA only require the reporting of daily averages or daily CEMS readings for normal operations. Detailed 1-hour data or usage by chamber should

[6] 59 FR 62585.

be maintained on-site and only submitted upon specific request or in the event of a deviation. We also suggest that the EPA adopt a "summary report" format for minor events. Consistent with other NESHAP standards, facilities should only be required to provide a detailed, line-item deviation report if the total duration of excess emissions or parameter exceedances exceeds 1% or CEMS downtime exceeds 5% of the total operating time during the reporting period. For facilities using PTEs, the requirement to provide monitoring data for each individual port or NDO within a PTE is physically and digitally overwhelming. Rather than requiring individual reporting for every sensor location, the EPA should allow for a "system-wide" status report. If the facility-wide pressure differential (*e.g.,* -0.007 inches of water) is maintained on a 24-hour average, the individual port data should be considered a component of internal recordkeeping rather than a mandatory reporting field. Shifting the reporting focus to long-term trends and corrective actions taken during "out-of-control" periods provides more value to regulators than the raw, high-frequency data currently proposed.

Question #11: Is PTE cost effective and feasible?

In response to the EPA's question 11, we provide the specific comments below.

We strongly support the EPA's proposal to remove the strict PTE mandate and the rigid application of EPA Method 204 for room air emissions. Method 204 was never designed for dynamic, high-traffic logistics environments where Natural Draft Openings (NDOs) and bay doors must routinely open for forklift traffic and product movement. However, the EPA's current proposal leaves a critical regulatory vacuum by failing to provide adequate, standardized federal guidance on how facilities should demonstrate capture efficiency in the absence of Method 204.

Without explicit federal guidance on how to determine Group 1 and Group 2 capture efficiency and DRE when Method 204 is not used, the EPA is effectively delegating this highly technical determination entirely to state and regional authorities. This will inevitably result in a patchwork of disjointed and conflicting compliance requirements across the country, as state agencies may implement arbitrary or overly restrictive mandates in the absence of a federal guidance. Rather than proposing exhaustive technical specifics that the EPA might further complicate, we urge the EPA to establish high-level regulatory guidelines and a standardized alternative capture efficiency protocol directly within the final rule to ensure national consistency.

We propose that the EPA adopt the *conceptual framework* of Method 204 to verify directional airflow and negative pressure, while explicitly uncoupling facilities from the method's rigid, geometric constraints. Specifically, the EPA should exclude requirements related to NDO distance from doors and eliminate prescriptive, overwhelming mandates regarding the number and location of Differential Pressure (DP) sensors. Instead of a "one-size-fits-all" geometric approach, the EPA should also loosely define the concept of maintaining negative pressure relative to the atmosphere (ensuring a cascade into Group 1 and Group 2 areas) without specifying the exact method of measurement. While we believe a third-party Method 204 demonstration is technically achievable for facilities designed for such standards, the method should be used as a verification tool rather than a rigid operational mandate.

While we support the EPA's proposed removal of the PTE mandate and the associated references to EPA Method 204, we recommend that the EPA retain PTE as an option to ensure both compliance and operational flexibility. Furthermore, we urge the EPA to replace prescriptive mandates with a requirement for a site-specific monitoring plan. This plan should serve as the primary vehicle for defining a facility's unique capture strategy. Specifically, we propose the following framework:

- **Eliminate Prescriptive Sensor Counts and Geometric Constraints:** We agree that the final rule should remove the fixed requirement for the number of (DP monitors and the restrictive NDO-to-VOC distance requirements. In addition, the final rule should allow facilities to outline the number and location of monitors within their individual monitoring plans based on unique facility airflow dynamics. The definition and maintenance of a PTE should be incorporated into this site-specific plan, allowing for a customized approach to demonstrating negative pressure relative to the atmosphere while maintaining a logical air cascade into Group 1 and Group 2 areas.

- **Implement High-Level Performance Alternatives:** For facilities where 100% capture via Method 204 is technically or architecturally infeasible, the EPA should provide clear guidance for a "contained facility" status. We propose two performance-based pathways in this regard:

    - Alternative 1: Maintain a negative pressure enclosure (*e.g.,* a minimum -0.007" of water) for Group 1 and Group 2 areas, coupled with a DRE of 80% or a reliable alternative outlet concentration limit (*e.g.,* < 1ppm) to account for low-inlet concentration scenarios; or

    - Alternative 2: A case-by-case demonstration of capture and control for Group 1 and Group 2 areas, as approved by the EPA, in instances where the geometric or pressure requirements of Alternative 1 remain prohibitive due to facility size or age.

By defining these high-level minimum requirements and shifting toward site-specific monitoring plans, and allowing a PTE as an option, the EPA can ensure that capture efficiency is verified through performance and data trends rather than rigid, "one-size-fits-all" geometry. This approach maintains environmental integrity while providing the necessary flexibility for facilities to operate within the realities of the modern medical device supply chain.

Question #12: Should the EPA not be allowed to treat unregulated sources of regulated pollutants the same as unregulated pollutants?

In response to the EPA's question 12, we provide the specific comments below.

We agree that the EPA should **not** treat unregulated sources (*i.e.,* newly identified emission points within a facility) of a regulated pollutant the same as entirely unregulated pollutants (*i.e.,* pollutants not previously regulated at all under Section 112).

The EPA rightly recognizes that the D.C. Circuit's decision in *Louisiana Environmental Action Network v. EPA (LEAN)*, 955 F.3d 1088 (D.C. Cir. 2020), does not require the EPA to regulate previously unregulated point sources of a previously regulated Hazardous Air Pollutant (HAP) like EtO. *LEAN* addressed a situation where, when revising emissions standards for a source category (pulp mill combustion sources) under section (d)(6), "EPA addressed some but not all the hazardous air pollutants they are known to emit." 955 F.3d at 1090. Those "other hazardous toxics" (*id*.)—different pollutants—had not been addressed in EPA's first round of standard setting in 2001. The EPA admitted that it was required to set Maximum Achievable Control Technology (MACT) floor standards under sections d(2)-(3) to control those different HAP (*id*. at 1091), yet did not do so before or when revising standards sixteen years later. *See also id.* at 1094 (noting petitioners had previously "urged the Agency to set limits for dozens of additional hazardous air pollutants that pulp mill combustion sources were known to emit," but the EPA had not done so). The D.C. Circuit held that, in that situation, the EPA was required to set standards for each of the different listed HAPs that the source category emits when revising standards for that category under section d(6). *Id.* at 1091; *see also id.* at 1096 (section (d)(6) requires the EPA to "include limits on *each hazardous air pollutant the category emits*") & 1100 (the EPA must "set limits on *the remaining hazardous air pollutants* emitted by pulp mill combustion sources") (emphasis added). The court did not hold—or even suggest—that the EPA must, when revising standards under d(6), set new standards for every previously unregulated point sources of a HAP, emissions of which from the broader source category are already regulated (as is the case here, where the EPA has already comprehensively regulated EtO emission from sterilizers).

Indeed, such a holding would be contrary to the D.C. Circuit's decision in *National Ass'n for Surface Finishing v. EPA*, 795 F.3d 1, 7-8 (D.C. Cir. 2015), where the court *rejected* the argument that the EPA must "calculate a new MACT floor" when it revises HAP standards under section (d)(6). The court noted that a "new MACT floor … would be based purely on the achievements of the best-performing facilities in the industry" after controls has been in place for years, and thus likely "more stringent than technology … revisions that take costs and other potentially constraining factors into existence." *Id.* at 8. The D.C. Circuit's description of such a hypothetical "new MACT floor" as predicated on the performance of "facilities" subject to prior HAP standards makes clear that the statute does not require the EPA to calculate new MACT floors for *every individual emission point* within a regulated facility when conducting technology review.

Further**,** from a technical and engineering standpoint, treating an unregulated individual emission point of a previously regulated HAP within a previously regulated source category the same as an entirely unregulated pollutant ignores the physical and engineering realities of commercial sterilization facilities and the vast differences in exhaust stream characteristics. In the commercial sterilization industry, the "unregulated sources" the EPA seeks to capture are predominantly fugitive room air emissions (*e.g.,* Group 1 and Group 2 areas). These emission streams possess entirely different aerodynamic and chemical profiles compared to the primary, regulated point sources within sterilization facilities (like the sterilization chamber exhaust).

- Point Sources vs. Fugitive Sources: Primary pollutant sources are highly concentrated, low-volume, and enclosed. Conversely, room air sources are highly dilute (often trace parts-per-billion levels) and high-volume.

- Control Device Limitations: Abatement technologies, particularly dry bed scrubbers, rely on concentration gradients. As discussed in our previous comments, it is technically impossible to achieve the same 99%+ Destruction and Removal Efficiency (DRE) on a dilute room air stream as it is on a concentrated chamber vent.

When the EPA treats these fugitive areas as if they were standalone new sources of entirely unregulated HAP, the Agency mistakenly attempts to apply point-source MACT floor mathematics to ambient environments. This results in mathematically unachievable emission limits that defy the physical limitations of commercial abatement equipment.

We strongly urge the EPA to respect the legal and technical distinctions between an entirely unregulated type of HAP emitted by a source category and a discrete emission point of a HAP already regulated under a NESHAP applicable to the source category. For minor or highly dilute emission points or streams (*e.g.,* room air and post-aeration warehousing), the EPA should not attempt to establish strict MACT floor percentage reductions. Instead, if the EPA determines that regulation of such additional emission points or streams within facilities is cost-effective and appropriate based on a technology review (which, as discussed above, the EPA may lawfully conclude is *not* the case), the EPA should rely on its authority under CAA Section 112(h) to establish practical, achievable work practice standards or alternative minimum outlet concentration limits that reflect the actual physical capabilities of modern control technology.

Question #13: Should the EPA remove PTE requirements?

In response to the EPA's question 13, we provide the specific comments below.

We strongly support the EPA's decision in the April 2, 2026, revised redline NESHAP to abandon the mandate for a PTE. As the EPA has rightly recognized, mandating a PTE is neither cost-effective nor broadly feasible for existing commercial sterilization facilities. Constructing a PTE is entirely cost-prohibitive for many existing facilities. The financial burden extends far beyond capital equipment; it includes extended facility shutdowns for installation, structural modifications, and exorbitant ongoing maintenance costs.

Commercial sterilizers have highly customized, site-specific footprints. Retrofitting an existing facility to meet EPA Method 204 PTE criteria is highly intrusive and, in some architectural configurations, physically impossible. This approach is also overly broad and short-sighted, as it fails to account for facilities that house multiple types of operations (*e.g.,* medical device packaging or general manufacturing) under the same roof. By imposing facility-wide enclosure requirements rather than focusing solely on the specific sterilization and aeration zones, the rule

inappropriately regulates the entire building. Furthermore, sweeping in large volumes of building air to maintain a PTE dilutes the EtO stream significantly. We have serious technical doubts regarding the feasibility of consistently meeting compliance requirements at these massive, ultra-low concentration airflows, without providing an alternative to the high DRE requirements.

While we support the EPA's removal of rigid PTE mandates and Method 204 references, this deletion creates a new challenge: ambiguity regarding exactly *how* a facility is expected to practically demonstrate an 80% DRE for room air and Group 2 emissions. If the final rule removes all federal methodological references without providing clear, flexible alternatives, facilities will be left vulnerable to conflicting interpretations by state and local permitting authorities, who may arbitrarily re-impose restrictive PTE mandates in their State Implementation Plans (SIPs).

As such, we urge the EPA to explicitly provide two distinct compliance pathways in the final rule to establish federal safe harbors and prevent state overreach:

- Method 204 Compliance Pathway: While we commend the EPA for removing the rigid, universal mandate and reference to Method 204 to allow for necessary operational flexibility, compliance via a PTE following Method 204 criteria should explicitly be retained as an option for facilities that can and choose to meet the pressure and geometric requirements as part of a site-specific performance plan. For facilities that choose to utilize this framework, it provides an established, streamlined method to confidently demonstrate 100% capture efficiency for compliance; or

- Mass-Balance Performance Pathway: Provide an equivalent alternative method (*e.g.,* Method 204E) to demonstrate actual capture efficiencies through mass balance or outlet testing for facilities where Method 204 geometry is prohibitive. This ensures that the rule regulates actual emission capture rather than building architecture.

Additionally, the EPA should include an explicit state program reciprocity provision so that if a facility demonstrates equivalent capture and control under a state-approved mass-balance or emissions cap, it satisfies the federal requirement.

In conclusion, we strongly support the EPA's proposal to remove the rigid PTE mandate. However, to ensure this intended flexibility is practically realized, the final rule must resolve the remaining technical bottlenecks within the regulatory text, while keeping a PTE as a viable option. By extending parametric averaging times to account for batch operations, retaining Method 204 and PS-19 strictly as options rather than mandates, and establishing clear federal alternative pathways for capture demonstration, the EPA can finalize a practical and effective rule that ensures robust emissions control without compromising the operational viability of the critical domestic medical sterilization supply chain.

Question #14: Should the EPA modify the definition of ''Operating day'' so that it applies to when an affected source is engaged in a sterilization operation, as opposed to the entire facility, as proposed?

In response to the EPA's question 14, we provide the specific comments below.

We support the EPA's proposal to modify the definition of "Operating Day" so that it applies specifically to periods when an affected source is actively engaged in a sterilization operation, rather than applying broadly to the entire facility. This is a significant and necessary improvement over the previous definition. However, the proposed modification remains incomplete and requires further clarification to be practically implementable.

While the EPA distinguishes between *new* and *existing* affected sources elsewhere in the rule, the modified definition of "Operating Day" fails to clearly delineate how it applies to specific, independent emission sources within a facility (*e.g.,* Sterilization Chamber Vents [SCVs] versus Aeration Room Vents [ARVs] versus Group 1/Group 2 Room Air Emissions). To be acceptable, the rule must explicitly define "affected source" within the context of this specific definition so that a facility does not trigger a facility-wide "Operating Day" simply because one isolated component is active.

Furthermore, we urge the EPA to explicitly address days where a facility is not actively sterilizing product, but where Room Air Emission systems are running continuously for safety and residual ventilation. If the definition of "Operating Day" encompasses these inactive periods, facilities will be forced to prove compliance on extremely dilute, baseline streams. Dry bed abatement systems and continuous monitors cannot mathematically or technically demonstrate a 80% DRE on trace-level room air streams due to inherently low inlet concentrations and the physical limits of measurement (Limit of Detection). Without excluding these "room-air only" days from the definition of an Operating Day, facilities will face unavoidable, false non-compliance events.

Question #15: Should the EPA modify the definitions of ''Sterilization operation'' and ''Single-item sterilization'' be modified as proposed?

In response to the EPA's question 15, we provide the specific comments below.

We have significant technical concerns regarding the proposed modifications to the definition of "Sterilization Operation" Specifically, the inclusion of "post-aeration operation" without clearly delineating it from final product storage under the umbrella of an active sterilization operation creates severe, unintended compliance hurdles.

The EPA's current proposed definition fails to adequately differentiate between the active handling of products immediately following formal aeration (true "post-aeration operations") and the purely passive storage of fully aerated products awaiting final distribution, including standard

facility shutdown modes. Categorizing these passive holding periods in a warehouse as an active "Sterilization operation" subjects them to the rule's most stringent monitoring and continuous compliance requirements.

During post-aeration operations, residual EtO off-gassing drops to ultra-trace, single-digit ppb levels. Because emission control systems (*e.g.,* scrubbers or dry beds) rely on mass transfer, their efficiency drops drastically when starved of target pollutant mass. Moreover, as previously noted in comments regarding CEMS, demonstrating compliance on these low-concentration streams is nearly impossible. Background noise, analyzer drift, and standard instrument limits of detection will routinely trigger false emission violations during these post-aeration phases. The EPA must sever passive post-aeration and product staging from the definition of an active "Sterilization Operation" to maintain a technologically feasible compliance standard.

We strongly recommend that the EPA remove the phrase "**When EtO is stored within the building**" from the regulatory definition of a "sterilization operation." Ethylene Oxide is delivered and stored in robust, heavily regulated, Department of Transportation (DOT)-approved pressurized cylinders and drums that act as completely sealed, passive containers prior to their connection to the sterilization chamber. Including passive storage within the definition of an active "Sterilization Operation" inappropriately broadens the regulatory scope, potentially forcing facilities to apply strict Group 1 capture and control requirements, *e.g.,* achieving an 80% emission reduction, to standard warehousing spaces where no active emissions are generated. Because it is mathematically impossible to demonstrate an 80% reduction on a passive storage room with near-zero baseline emissions, the definition should be strictly limited to the active dispensing, processing, and aeration areas where EtO is actively introduced and managed.

Question #16: Should the EPA modify the definition of ''Indoor EtO storage'' to clarify that EtO storage media of calibration gas cylinders that is only used to verify the performance of EtO CEMS is excluded, as proposed?

In response to the EPA's question 16, we provide the specific comments below.

We strongly support the EPA's proposal to modify the definition of "Indoor EtO Storage" to explicitly exclude EtO calibration gas cylinders. However, we urge the EPA to expand this exclusion to encompass all EtO calibration gases maintained on-site, rather than limiting the exemption strictly to gases used to verify the performance of EtO CEMS. Furthermore, as established in our previous comments, we reiterate our recommendation that passive EtO storage should be removed entirely from the overarching definition of a "Sterilization Operation". These modifications are practical, necessary, and effectively resolve conflicting regulatory burdens.

Under EPA Performance Specification 19 (PS-19), facilities utilizing CEMS are mandated to perform rigorous, daily calibration drift checks. To meet this federal requirement, facilities must maintain a continuous supply of trace-level EtO calibration gases on-site. In addition, other EtO detection systems not used for Subpart O compliance such as gas chromatographs or workplace monitoring systems may be used in sterilization facilities and require EtO calibration gases. These

cylinders are fundamentally different in both volume and concentration (often containing highly dilute, ppm- or ppb-level EtO) from the EtO drums used to supply sterilization operations. Subjecting these critical compliance tools to the stringent emission control and monitoring requirements designed for "Indoor EtO Storage" would create an illogical regulatory loop—essentially penalizing facilities for maintaining the very equipment required to prove compliance with the NESHAP.

Question #17: Should the EPA modify 40 CFR 63.363(b)(1) to indicate that CEMS measurement points do not need to be equidistant so long as the response times are similar, as proposed?

In response to the EPA's question 17, we provide the specific comments below.

We support the EPA's proposal to modify 40 CFR 63.363(b)(1) to clarify that CEMS measurement points do not need to be physically equidistant from the analyzer, provided that the system response times are similar.

This proposed modification reflects the practical and physical realities of facility layouts. In many sterilization facilities, it is structurally impossible or highly impractical to route sample lines so that every measurement point is exactly equidistant from a centralized CEMS analyzer. Mandating equidistant lines places an unnecessary engineering constraint on facilities without yielding any additional environmental benefit. Furthermore, analytical best practices dictate that sample lines should always be kept as short as possible. If a facility were forced to maintain equal distances, it could have adverse effect to artificially and unnecessarily result in facilities extending their shorter sample lines to match the length of the furthest measurement point. This forced extension not only increases the capital costs associated with expensive heated sample lines, but more importantly, it introduces additional internal surface area to the system. For a highly reactive compound like EtO, increased surface area drastically raises the potential for the sample to adsorb onto the tubing walls before reaching the analyzer, thereby compromising the integrity and accuracy of the measurement.

The primary regulatory objective is to ensure accurate, timely, and comparable data collection. As the EPA correctly notes in the proposal, this objective is achieved through consistent response times, not physical line length. Facilities can easily and effectively compensate for varying sample line lengths through standard engineering practices, such as:

- Utilizing different sampling pumps sized appropriately for the specific line length;
- Adjusting individual pump flow rates to equalize the transit time of the sample to the analyzer; or
- Staggering sample extraction times to ensure cohesive and synchronized data analysis.

By finalizing this modification, the EPA will provide facilities with the necessary design flexibility to implement effective CEMS architectures while minimizing sample line loss and maintaining the stringent data quality and timing standards required for compliance demonstration. We urge the EPA to finalize this practical update as proposed.

Question #18: Should the EPA remove all references to 40 CFR part 75 for flow rate monitors within Appendix A to the Commercial Sterilization Facilities NESHAP and should we make the corrections described in table 6 be made, as proposed?

In response to EPA's question 18, we provide the specific comments below.

We recognize that 40 CFR Part 75 is a separate and highly detailed regulatory framework governing CEMS. It contains stringent, specific procedures for the initial certification, continuous verification, and quality assurance/quality control (QA/QC) of monitors. Because Part 75 was originally developed for large combustion sources (*e.g.,* those under the Acid Rain Program), its rigorous verification procedures are often overly complex or poorly tailored to the operational realities, scale, and exhaust profiles of commercial EtO sterilization facilities.

We generally support removing overly burdensome or misaligned regulatory cross-references, but we also need to ensure that the replacement protocols (or the remaining Part 60/63 requirements) provide a clear, practical, and achievable pathway. Specifically, we note that the requirements for continuous flow monitoring systems are proposed in the April 2, 2026, redline rule version in 40 CFR § 63.364(g) and would apply to both CEMS and parametric flue gas monitoring applications if finalized. Originally, these flow monitoring requirements were limited to flow monitoring as an option to demonstrate compliance with PTE. Requirements developed specifically for PTE compliance are not directly transferable to the broader range of flow monitoring applications now proposed in the rule, including CEMS and flue gas monitoring. While CEMS relies on flow monitoring to calculate emissions, flow monitoring associated with emission control equipment serves a fundamentally different purpose. In that context, flow monitoring functions solely as a monitoring parameter to demonstrate that gas flow does not exceed the designed range of the emission control unit. However, design flow rates are typically used to purchase appropriately rated equipment (*i.e.* fans) for emission control devices. As stated previously, we strongly recommend the EPA simplify the OPL mandates focusing on the key monitoring parameters which would remove the need to monitor flow. If flow is to remain as a required parameter, the distinct compliance objectives justify differing levels of rigor in flow measurements. Accordingly, the proposed procedures should be carefully tailored to reflect these differing uses. The flow monitoring requirements that mandate extensive QA/QC and verification procedures may be appropriate for flow monitors associated with CEMS but are excessive for parametric monitoring. The EPA should therefore reconsider the scope and rigor of the proposed flow monitoring requirements based on their intended compliance purpose—specifically distinguishing between direct emissions monitoring and parametric monitoring.

Question #19: Should the EPA add the phrase ''RATA test runs must be at least 21 minutes in length'' to the end of section 2.2.4.3 of Appendix A to subpart O, as proposed?

In response to the EPA's question 19, we provide the specific comments below.

While we understand the EPA's intent to standardize Relative Accuracy Test Audit (RATA) procedures, especially in light of the proposed removal of 40 CFR Part 75 references (as discussed in Question #18), mandating a strict 21-minute minimum run length in Appendix A could inadvertently limit future operational and technological flexibility. We acknowledge that a 21-minute minimum run time is generally consistent with current, reasonable industry practices; however, explicitly mandating this duration within the rule text creates a rigid chronological constraint that ignores the realities of modern data acquisition and established Performance Specifications.

Historically, the 21-minute RATA run time was established to ensure the collection of at least 20 valid one-minute data points, which is the statistical minimum required to calculate a meaningful average and assess variability without the results being overly skewed by outliers. However, modern Data Acquisition and Handling Systems (DAHS) for both facility CEMS and Reference Method (RM) CEMS can record data at much higher frequencies (*e.g.,* 10-second averages). As analytical technology advances, facilities can achieve the required volume of data points for stable statistics in a shorter chronological window. Locking the rule into a 21-minute minimum ignores this technological evolution in data density.

RATA run times and data collection requirements are typically dictated by EPA Performance Specifications (*e.g.,* PS-19), not the underlying Reference Methods (which typically dictate sample *volume*, rather than time). Hardcoding a 21-minute temporal requirement directly into Subpart O is redundant and risks creating conflicting mandates if future Performance Specifications are updated to accommodate new statistical approaches or sampling paradigms.

Many facilities operate under state-level delegated authority, where environmental agencies have their own established methodologies and test plan requirements for conducting RATAs. Leaving the run time flexible to align with the applicable Performance Specifications helps avoid potential friction between a codified federal time mandate and site-specific state test plans, streamlining the approval process for all parties.

Instead of prescribing a 21-minute minimum run length in Section 2.2.4.3, we recommend that the EPA allow the RATA test run duration to default to the specific requirements of the applicable Performance Specification and the facility's approved, site-specific test plan.

Question #20: Should the EPA retain the compliance deadlines promulgated in the 2024 Final Rule for the section 112(d) standards as proposed?

In response to the EPA's question 20, we provide the specific comments below.

We strongly urge the EPA **not** to retain the strict compliance deadlines promulgated in the 2024 Final Rule for the Section 112(d) technology-based standards.

To begin, we do not think it is appropriate to retain compliance deadlines from an over two-year-old rule that has long been under reconsideration and its application to most regulated facilities deferred by Presidential Proclamation. Regulated sources of EtO have been relying on the ongoing reconsideration proceeding and the Presidential compliance exemptions, and accordingly have not taken the many steps necessary to come into compliance by the original deadlines. To now reimpose those original deadlines—which, for Section 112(d) standards, would require compliance in April 2027—would unfairly disrupt regulated facilities' expectations. Further, even if the EPA finalizes the rescission of the stringent risk-based standards set in the 2024 Final Rule, retaining the prior deadline for new or revised section 112(d) standards would place sterilization facilities in an untenable position given the limited time remaining to acquire, install, and successfully operate the control technology needed to comply. We accordingly ask that the EPA instead adhere to its usual approach of establishing new compliance deadlines in the forthcoming new final rule based on the publication date, and we specifically propose that the EPA provide up to three years from that new final rule for regulated sources to come into compliance with the standards set or revised therein.

Allotting up to three years for compliance from publication of the new final rule is also necessary as a practical matter. From a technical, engineering, and logistical perspective, the original 2- to 3-year timelines established in April 2024 are insufficient for commercial sterilization facilities to achieve full, safe compliance without risking severe disruptions to facility operations and the sterile medical supply chain. While we recognize the EPA's mandate to implement technology-based standards, retaining the 2024 deadlines ignores the practical realities of industrial retrofitting. We highlight the following specific technical barriers that necessitate setting a new, more extended deadline.

The NESHAP requirements trigger simultaneous, industry-wide demand for highly specialized air pollution control devices (*e.g.,* multi-stage dry bed scrubbers, catalytic oxidizers) and advanced CEMS like Cavity Ring-Down Spectroscopy (CRDS) and Optically Enhanced FTIR. There is a highly limited pool of vendors capable of manufacturing and validating this equipment. Current lead times for procurement and delivery alone frequently exceed 12 to 18 months, leaving almost no time within the original 2024 deadline for installation, testing, and troubleshooting.

Complying with the Section 112(d) standards often requires fundamental changes to a facility's HVAC architecture to meet advanced room air capture and routing requirements. This is not a "plug-and-play" process. It requires:

- Extensive computational fluid dynamics (CFD) modeling;
- Structural modifications to existing buildings to route new, heavy ductwork; and

- Significant facility downtime. Rushing the engineering phase of aerodynamic capture systems risks creating pressure imbalances that could compromise both emission capture efficiency and occupational safety on the facility floor.

Also, before any physical construction or installation of new control devices can commence, facilities must draft, submit, and receive approval for state or local air permit modifications (*e.g.,* Title V or New Source Review permits). Nationwide, environmental regulatory agencies are experiencing significant backlogs. It is common for permit approvals to take 9 to 12 months. This administrative reality effectively consumes up to half of the EPA's original compliance window before ground can even be broken on physical improvements.

Once equipment is physically installed, facilities must contract highly specialized, third-party stack testing firms to conduct initial performance tests and CEMS Relative Accuracy Test Audits (RATAs). Because the 2024 rule forces dozens of facilities nationwide to comply simultaneously, the limited pool of qualified testing firms possessing the necessary sub-ppb detection equipment (*e.g.,* EPA Method 320 capabilities) has created a severe scheduling bottleneck.

Based on these compounding technical factors, we recommend that the EPA formally extend the compliance deadlines for the Section 112(d) standards by three years. This extension is necessary to accommodate standard industrial construction timelines, equipment procurement realities, and local permitting requirements, ensuring that facilities can successfully implement robust compliance measures without halting the sterilization of critical medical devices.

Question #21: Would the proposed standards result in a similar risk of capacity constraints as the 2024 Final Rule in the sterilization sector?

In response to the EPA's question 21, we provide the specific comments below.

The threat to sterilization capacity posed by the 2024 Final Rule was the result of extremely stringent risk-based standards (which the EPA now rightly proposes to rescind) and prescriptive operational mandates (*e.g.,* mandatory CEMS monitoring requirements) that could have forced facility shutdowns or prevent the processing of essential medical devices due to rigid constraints and requirements. The EPA should be aware that due to the stringent requirements, some sterilizers made the decision to cease operation, further constraining the industry's capacity to sterilize sufficient medical products to serve the U.S. market.

If the EPA finalizes the reconsideration rule to (1) eliminate the legally impermissible and technically unachievable risk-based standards, (2) allow sources to choose between CEMS and parametric monitoring to demonstrate compliance, (3) remove mandatory PTE requirements, (4) revise the technology-based standards along the lines proposed, and (5) provide for alternative outlet concentrations alongside those percentage reduction-based standards, we anticipate that sterilizers will be able to continue operations at a level that maintains the industry's ability to meet U.S. demand for sterilized medical devices and other essential products.

Other Comment #22 : EPA should allow for an alternative outlet concentration limit

We provide specific comments below to urge the EPA to allow for an alternative outlet concentration limit in the Sterilizer NESHAP for all emission sources.

While we appreciate the EPA's efforts in the proposed reconsideration to address the technical challenges of the Commercial Sterilization Facilities NESHAP, we strongly urge the Agency to include an alternative outlet concentration limit (*e.g.,* <1 ppm threshold) alongside the respective percentage reduction standards for all emission sources as proposed in the final rule.

Basing compliance strictly on a high percentage reduction creates severe, unintended mathematical and analytical compliance traps for facilities, particularly during the later stages of the sterilization process or Group 1 and/or Group 2 room air emissions, where the EtO inlet concentration is very low. Furthermore, because the EPA's expansive definition of "Sterilization Operation" encompasses post-aeration processes characterized by very low inlet concentrations, demonstrating a strict percentage-based removal efficiency is often mathematically infeasible, also the emission control systems are unable to achieve the required DREs at these very low inlet concentrations, underscoring the critical need for an alternative outlet concentration standard. As such, an alternative outlet concentration limit is technically sound, environmentally protective, and aligns with established EPA precedent in other NESHAP standards.

Commercial EtO sterilization is a batch process, meaning the concentration of EtO routed to emission control devices is not constant. During the initial evacuation of a sterilization chamber, the inlet concentration is high, and achieving a 99.6% destruction efficiency is mathematically straightforward. However, during the aeration phase or chamber back-venting, or for the Group 1 and/or Group 2 room air emissions, the inlet EtO concentration drops precipitously. As the inlet concentration (the denominator in a DRE calculation) approaches zero, it becomes mathematically impossible to demonstrate an 80% let alone a 99.6% reduction. Even if the control device is operating at peak efficiency and the absolute mass of EtO being emitted is virtually zero, a facility will technically "fail" the performance test because the mathematical ratio breaks down at low inlet loadings. For example, an inlet concentration of 0.001 ppm in Group 1 or Group 2 areas requires the detection system to reliably measure a 0.2 ppb EtO concentration at the outlet to achieve an 80% DRE, and this 0.2 ppb level is below the "representative detection level" by all detection systems.

As commented previously, when the inlet concentration is very low, the required outlet concentration to prove a 99.6% or 80% DRE frequently falls below the reliable detection threshold of even advanced FTIR or CRDS systems. At these trace levels, standard instrument "noise" or minor baseline drift is recorded as a small, positive EtO value. When this trace artificial value (the numerator) is divided by the exceptionally low inlet value (the denominator), the resulting calculated DRE can easily drop artificially below the 80% requirement for Group 1 and Group 2 emissions, falsely indicating non-compliance.

The EPA has long recognized the "low inlet concentration" dilemma in other NESHAP rulemakings. For example, in the New Source Performance Standards for the Synthetic Organic Chemical Manufacturing Industry and National Emission Standards for Hazardous Air Pollutants for the Synthetic Organic Miscellaneous Organic Chemical Manufacturing Industry and Group I & II Polymers and Resins Industry (HON Rule) finalized on May 16, 2024,[7] and the National Emission Standards for Hazardous Air Pollutants: Polyether Polyols Production Industry Review (PEPO NESHAP) published on March 18, 2026,[8] the EPA includes a dual standard: requiring a 98-99% weight reduction OR an outlet concentration of < 20 ppm.

We strongly recommend that the EPA amend the final reconsidered NESHAP to allow facilities the option to comply by either achieving the specified percentage reduction (*e.g.,* 99.6% for ARV or 80% for G1 and G2), OR maintaining an alternative outlet concentration limit (*e.g.,* < 1 ppm) for all the emission streams.

Other Comment #23 : Group 2 definition

We provide specific comments below to urge the EPA to modify and clarify the Group 2 definition and its associated compliance pathways within the Sterilizer NESHAP.

We support the EPA's proposal for existing Group 2 Room Air emissions that allows facilities to choose to either lower the EtO concentration to 1 ppm before opening the aeration chamber, or achieve an 80 percent reduction of EtO (80% DRE). The EPA's April 2, 2026, revised redline NESHAP further specifies that facilities may choose a parametric approach... or a direct measurement approach for determining the sterilization chamber end-cycle EtO concentration. While industrial chambers are not typically able to achieve levels of 1 ppm at the end of the cycle, this option for Group 2, in addition to the 80% DRE and the alternative outlet concentration limit (< 1 ppm) that we strongly advocate for, provides much-needed flexibility for facilities of various sizes and designs to meet compliance requirements.

Fundamentally, because Group 2 areas deal exclusively with post-aeration product, the primary source of any potential fugitive EtO is the residual off-gassing from these sterilized materials. If a facility can successfully demonstrate that the EtO concentration within the chamber or connected process has been reduced to < 1 ppm before the product is even removed, or if it achieves an outlet concentration of less than 1 ppm, the residual off-gassing potential in subsequent Group 2 areas becomes effectively negligible. Mandating an 80% capture and control requirement on massive post-aeration warehouses and manufacturing or packaging operations when the source product has already been verified to emit negligible EtO is technically redundant and imposes severe, unnecessary capital costs. As such, it is more practical to keep the 80% DRE as an option instead of a mandate.

[7] 89 FR 42982, see page 42935

[8] 91 FR 13116, see page 13156

Despite this proposed flexibility, the current definition of "Group 2 Room Air" (which states that "Group 2 room air emissions mean emissions from post-aeration handling of sterilized material") and its associated spatial and operational boundaries remain problematically vague. Specifically, it is highly unclear where a "Group 2" zone begins and ends. Facilities are left without clear guidance on whether auxiliary spaces, administrative zones, loading docks, or long-term finished-goods warehousing are legally captured under the "Group 2" umbrella. To resolve this, the EPA must explicitly define the boundaries of Group 2 operations. Without such clarification, it is structurally impossible for facilities to engineer compliant negative-pressure capture systems, accurately size dry bed abatement systems, or determine the correct inlet mass required to calculate an accurate DRE.

Other Comment #24 : Group 1 and Group 2 DRE feasibility

We provide specific comments below to urge the EPA to reconsider Group 1 and Group 2 DRE feasibility in the Sterilizer NESHAP.

We have significant concerns regarding the technical feasibility of demonstrating strict DRE requirements for Group 1 and Group 2 room air emissions. While we support the EPA's goal of reducing fugitive emissions, the aerodynamic and analytical realities of treating highly dilute room air make a strict DRE mandate unworkable.

When a facility utilizes the SWEL 1 framework, compliance effectively becomes a mass-based calculation (*e.g.,* pounds per year or pounds per hour). However, if a facility does not or cannot use the SWEL approach, it is forced to measure and calculate a strict DRE percentage for Group 1 and Group 2 room air controls. Because the inlet concentration (the denominator in the DRE calculation) is already exceptionally low, proving an 80% or greater reduction pushes the required outlet concentration to mathematically impossible levels, often falling below the Method Detection Limits (MDL) that even the most advanced CEMS or FTIR analyzers can reliably measure. For example, an inlet concentration of 0.001 ppm in Group 1 or Group 2 areas requires the detection system to reliably measure a 0.2 ppb EtO concentration at the outlet to achieve an 80% DRE, and this 0.2 ppb level is below the "representative detection level" of virtually all detection systems.

As such, we strongly advocate that the EPA provide an alternative minimum outlet concentration limit for Group 1 and Group 2 room air emission controls and other emission sources. The mathematical calculation, performance guarantees from emission control manufacturers, and precedence set in other recently finalized NESHAP rules strongly support this request.

Other Comment #25 : CEMS and PS19 requirements

We provide specific comments below to urge the EPA to modify CEMS and PS19 requirements in the Sterilizer NESHAP.

We strongly support the EPA's proposed 2026 reconsideration to remove the strict, mandatory CEMS requirement for facilities using more than 10 tons of Ethylene Oxide (EtO) per year. Opting for a flexible framework that allows facilities to choose continuous parametric monitoring as a permanent alternative to CEMS is a highly practical and effective approach to ensuring compliance.

Our support is grounded in the reality that the rigid application of PS-19 presents severe technical, logistical, and mathematical hurdles. Making CEMS and PS-19 a universal mandate is operationally infeasible for the commercial sterilization industry for several key reasons:

- PS-19 imposes stringent Daily Calibration Drift (CD) limits and ongoing Quality Assurance procedures, such as Cylinder Gas Audits (CGA) or Dynamic Spiking (DS). At the ultra-low measurement ranges required for EtO, standard instrument baseline noise can easily mathematically exceed PS-19's allowable percentage drift limits. This rigidity triggers false out-of-control periods and invalidates data, even when the analyzer is functioning perfectly.

- PS-19 was originally designed for continuous, steady-state chemical manufacturing. Commercial sterilization is a dynamic batch process characterized by rapid operational shifts (evacuation, back-venting, and aeration). The routine spikes, lulls, and wildly fluctuating concentrations inherent to batch processing make it incredibly difficult to meet the strict cycle-time and steady-state validation requirements of PS-19, causing continuous data management and invalidation headaches.

- PS-19 requires initial and annual Relative Accuracy Test Audits (RATAs) utilizing EPA Method 320. The nationwide pool of third-party stack testing firms equipped and accredited to perform sub-ppb Method 320 RATAs is far too small to support the entire sterilization industry simultaneously. A mandate would inevitably force facilities into non-compliance due to third-party supply chain and scheduling bottlenecks that are entirely outside their control.

Because parametric monitoring provides a highly reliable, proven, and continuous demonstration of control device efficacy without the fragility of CEMS, we strongly recommend that the EPA codify CEMS and by extension compliance with PS-19, strictly as an optional compliance pathway, not a mandate. Retaining PS-19 as an option provides flexibility for facilities that wish to adopt advanced analyzer technologies, while allowing the broader industry to maintain reliable, continuous compliance through standard parametric monitoring.