# Preliminary Comments from Members of the CASAC on EPA's *Integrated Science Assessment for Ozone and Related Photochemical Oxidants (External Review Draft – September 2019)*

**Received as of 11-27-19**

**Dr. Tony Cox** ..... **2**
**Dr. Mark Frampton** ..... **38**
**Dr. Sabine Lange** ..... **45**
**Dr. Ronald J. Kendall** ..... **83**
**Dr. Corey Masuca** ..... **100**

# Dr. Tony Cox

## Responses to Charge Questions

*The Executive Summary is intended to provide a concise synopsis of the key findings and conclusions of the Ozone ISA for a broad range of audiences.* ***Please comment on the clarity*** *with which the Executive Summary communicates the key information from the draft ISA.* ***Please provide recommendations on information that should be added*** *or information that should be left for discussion in the Integrated Synthesis and accompanying appendices of the draft ISA.*

The key information provided in the draft ISA and its Executive Summary is unclear. Concerns about lack of clarity in how key results are derived, expressed, and communicated have been raised in numerous public comments in this and previous NAAQS review cycles. They are still not addressed in the current draft ISA.

In addition to inviting public comments, this review cycle for the first time gave a panel of external expert consultants an opportunity to comment directly on the following question: *Is the scientific information provided by the ISA clear?* Appendix A provides responses from the consultants to this and other questions. Their main answer to this question is no, for at least the following reasons:

- *Criteria for selecting and weighting studies, and how key conclusions are derived from them, are not clear.*
- *It is unclear how, if at all, conclusions would change if consistent criteria were systematically applied* for selecting, evaluating, summarizing, and synthesizing studies.
- *The draft ISA and its Executive Summary do not provide comprehensive quantitative uncertainty and sensitivity analyses* showing how conclusions change for plausible variations in assumptions, interpretations of undefined and vague terms, selection and weighting of studies, and judgments on which the conclusions depend.
- *Causal determination judgments appear to be ambiguous, subjective, and sometimes arbitrary.* Several external consultants commented that different people might well make the determinations in very different ways from the same data. The evidence presented often does not seem to clearly support one causal determination to the exclusion of others. These experts noted that they could not guess, for any particular body of evidence, which causal determination category EPA will choose to describe it. By this criterion, the causal determinations do not seem to follow clearly from the evidence presented, but incorporate an arbitrary (unpredictable) element. The draft ISA provides no clear

objective basis for determining or predicting from facts and data which causal determination (if any) is right.

- *What the causal determinations mean, and what they imply for empirical observations, is unclear*. For example multiple external experts agreed that the term "causal" is used in the draft ISA without distinguishing among importantly different causal concepts. No distinction is made among necessary, sufficient, INUS, and other forms of causation. Yet, a stated conclusion such as a determination that a specific C-R association is "causal" or "likely to be causal," is often correct for some of these causal concepts and incorrect for others. Thus, the use of vague general terms such as "causal" or "likely to be causal" to communicate key conclusions in the draft ISA makes their meanings highly ambiguous and impossible to determine with clarity. As stated by Dr. Rhomberg, "If patterns of association are plausibly explained by underlying causation, this is taken [in the draft ISA] as sufficient evidence for such causation (when one should actually be comparing the hypothesized causative actions against other competing explanations for the patterns), and if such causation is inferred, it is taken to be universal, applying to other settings on the usually poorly stated (much less justified) presumption that the causation is universal and largely independent of other circumstances." Such inferences and presumptions, which pervade the draft ISA, are scientifically unsound, making the validity as well as the meanings of key conclusions unclear. Leaving unspecified exactly what the draft ISA means (e.g., one of these specific concepts, or perhaps something else) when it uses the term "causal" therefore makes it impossible to determine whether its key conclusions expressed using this term are correct, or even what they are intended to mean. This situation is sometimes referred to in other areas of science as "not even wrong," i.e., key findings depend on arbitrary-seeming judgments expressed using poorly defined, vague, or ambiguous terms, so that it is not clear what they mean (or how they could be tested or falsified by data), let alone whether they are correct.
- *The draft ISA's treatment of wildfire contributions to ozone exposure,* and their implications for NAAQS, are unclear.
- *It is unclear to what extent the ozone-associated physiological effects discussed in the draft ISA are transient* and to what extent they represent persistent, avoidable harms that could be reduced by further reducing ozone levels. Yet, this is crucial information for policy makers.

Information that should be added to the Executive Summary includes the following:

1. *Discussion of how changes in public health effects depend on changes in ozone levels.* This is the most important scientific topic for informing the PA. It is not addressed in the

draft ISA. A useful ISA should address the extent to which reducing ozone has been found to cause reductions in public health risks and improvements in public health and welfare, and the extent to which additional reduction should or should not be expected to cause further benefits. It should quantify uncertainties about the answers. As stated by Dr. North, we "should be seeking to evaluate manipulative or interventional causation, that is, determining how many people might be added or subtracted from having their health protected with an adequate margin of safety by a change in the primary NAAQS standard." The draft ISA does not present relevant scientific information to use in addressing this question. Dr. Lipfert notes that "The ultimate test of causality is whether health has actually improved since the late 1970s in response to peak O3 levels reduced by a factor of 5 in conjunction with coincident trends in spatial patterns of reduced smoking and improved medical care. A search of PubMed found no support for such improvement." Rather than relying on searches by external experts, it would be far preferable for the EPA itself to address in the final ISA the key policy-relevant scientific question of how changes in public health risks depend on changes in ozone levels.

2. *Summary of results from a systematic review and critical evaluation and synthesis of relevant studies, including negative ones that have been omitted from the draft ISA*. The comments in Appendix A and other comments received from the external expert consultants and the public identify some of the omitted relevant studies (e.g., Moore et al. 2013).
3. *Detailed discussion of possible confounding*, and how it was or was not addressed in reaching causal conclusions. As noted by Dr. North in the context of answering a specific question, "I perceive that the kind of discussion needed on confounding was not present in the ISA, just a judgment of 'likely to be causal.' And I do not find such judgments useful in the absence of detailed discussion of possible confounding."
4. *Results of systematic evaluations of study quality*, *using consistently applied criteria,* showing how each key study included performs on each specific quality criterion relevant for evaluating individual studies and drawing valid causal conclusions (e.g., identification of appropriate adjustment sets, control of observed confounders, control of residual confounding, control of latent confounding, quantification of exposure estimation errors and uncertainties, adjustment of effects estimates for errors-in-variables, quantification of model uncertainty, adjustment of reported significance levels and confidence levels for model uncertainty, control for multiple testing bias, use of appropriate control groups, tests for internal validity, tests for external validity and invariant causal prediction property). A matrix (possibly color-coded, as suggested by Dr. Goodman in public comments on PM2.5) summarizing these results could provide great insight into the state

of the literature and the strengths and limitations of individual key studies used by the EPA in reaching its conclusions.

5. *Discussion of causal biological mechanisms of inflammation-related health effects* and their implications for biologically realistic causal C-R functions. For example, the ISA should discuss recent evidence on mechanisms and modes of action (possibly including the role of the NLRP3 inflammasome in inflammation-mediated responses to ozone exposures) and their implications for the shapes of causal C-R functions describing health responses to changes in exposures below the current NAAQS. (Importantly, such causal C-R functions should not be confused with regression C-R functions.)
6. *Results of comprehensive, quantitative uncertainty and sensitivity analyses* showing how conclusions change for variations in inputs, including selection and weighting of studies, alternative interpretations of study results, corrections for confounding, corrections for measurement errors, corrections for historical trends, modeling choices and assumptions, interpretations of undefined and vague terms, and subjective judgments and unverified assumptions on which conclusions depend.

*The Integrated Synthesis presents and synthesizes the overall conclusions from the subsequent detailed appendices of the draft ISA and characterizes available scientific information on policy relevant issues. Please comment on the* ***usefulness and effectiveness of the summary*** *presentation. Please provide* ***recommendations on approaches that may improve the communication of key findings*** *to varied audiences and the* ***synthesis of available information*** *across subject areas.* ***What information should be added*** *or is more appropriate to leave for discussion in the subsequent detailed appendices?*

The Integrated Synthesis has the following limitations that limit its usefulness and effectiveness.

- *Biased selection of studies.* Multiple readers of the draft ISA, including several external expert consultants, have noted that it omits many relevant studies, especially those that do not conclude that ozone is associated with adverse health effects. This undermines the credibility, completeness, and scientific usefulness and effectiveness of the draft ISA, especially to the extent that it creates an impression that the evidence presented has been selected to support a narrative rather than to neutrally convey the current state of the art in the underlying science.
- *Literature on nonlinear effects is not well covered.* The draft ISA does not adequately cover the recent scientific literature on nonlinear C-R functions for ozone. For example, p. IS-88 states that "Examination of the concentration-response (C-R) relationship has primarily been conducted in studies of short-term ozone exposure and respiratory health effects or mortality, with some more recent studies characterizing this relationship for

long-term ozone exposure and mortality. Across recent studies that used a variety of statistical methods to examine potential deviations from linearity, **evidence continues to support a linear C-R relationship**, but with less certainty in the shape of the curve at lower concentrations (i.e., below 30−40 ppb)." This contrasts with a substantial literature, disregarded in the draft ISA, on nonlinear C-R relationships. For example,

- Bae et al. (2015) report that "The mean O3 concentration did not differ greatly between Korea and Japan, which were 26.2 ppb and 24.2 ppb, respectively. Seven out of 13 cities showed better fits for the spline model compared with the linear model, supporting a **non-linear relationships between O3 concentration and mortality**. **All of the 7 cities showed J or U shaped associations suggesting the existence of thresholds**. The range of city-specific thresholds was from 11 to 34 ppb. **The city-combined analysis also showed a non-linear association with a threshold around 30-40 ppb**." (Bae S, Lim YH, Kashima S, Yorifuji T, Honda Y, Kim H, Hong YC. Non-Linear Concentration-Response Relationships between Ambient Ozone and Daily Mortality.) PLoS One. 2015 Jun 15;10(6):e0129423. doi: 10.1371/journal.pone.0129423.
- Seltzer et al. (2018) state that "Long-term ozone (O3) exposure estimates from chemical transport models are frequently paired with exposure-response relationships from epidemiological studies to estimate associated health burdens. **Impact estimates using such methods can include biases from model-derived exposure estimates**. We use data solely from dense ground-based monitoring networks in the United States, Europe, and China for 2015 to estimate long-term O3 exposure and calculate premature respiratory mortality using exposure-response relationships derived from two separate analyses of the American Cancer Society Cancer Prevention Study-II(ACS CPS-II) cohort. …**Both sets of results are lower (~20%–60%) on a region-by region basis than analogous prior studies based solely on modeled O3, due in large part to the fact that the latter tends to be high biased in estimating exposure.** This study highlights the utility of dense observation networks in estimating exposure to long-term O3 exposure and provides an observational constraint on subsequent health burdens for three regions of the world. In addition, **these results demonstrate how small biases in modeled results of long-term O3 exposure can amplify estimated health impacts due to nonlinear exposure-response curves**."
- Wilson et al. (2014) report that, even in modeling that constrains ozone C-R functions for mortality to be monotonic (disallowing J-shaped or U-shaped relations such as those reported by Bae et al.), "We then examine the synergistic

effect of ozone and temperature both nationally and locally and **find evidence of a nonlinear ozone effect** and an ozone-temperature interaction at higher temperatures and ozone concentrations." The draft ISA reports the nonlinear interaction from this study (p. 6-12), but does not mention the "evidence of a nonlinear ozone effect."

That the draft ISA does not mention such results on nonlinear C-R functions for ozone suggests that readers interested in understanding the available scientific evidence on ozone C-R functions must do their own research: the draft ISA fails to cover many studies and results that disagree with its narrative (in this case, that "evidence continues to support a linear C-R relationship"). This lack of coverage of diverse findings in the literature undermines the credibility, effectiveness and usefulness of the draft ISA.

- *Summaries of relevant literature are incomplete and of questionable accuracy.* The draft ISA does not provide a comprehensive or trustworthy summary of available scientific evidence, even for studies and authors that it cites. For example:
    - Page 3-91 of the ISA states that "A limited number of recent studies provide evidence of an association between long-term exposure to ozone and asthma development in children. … An overview of the evidence is provided below. A recent CHS analysis examined asthma incidence in relation to improved air quality in nine southern California communities (Garcia et al., 2019). ***Decreases in baseline ozone concentrations in three CHS cohorts, enrolled in 1993, 1996, and 2006, were associated with decreased asthma incidence.***" However, Garcia et al. (2019) actually state that "Among children in Southern California, decreases in ambient nitrogen dioxide and PM2.5 between 1993 and 2014 were significantly associated with lower asthma incidence. ***There were no statistically significant associations for ozone*** or PM10." (Garcia E, Berhane KT, Islam T, McConnell R, Urman R, Chen Z, Gilliland FD. Association of Changes in Air Quality With Incident Asthma in Children in California, 1993-2014. JAMA. 2019 May 21;321(19):1906-1915. doi: 10.1001/jama.2019.5357. Emphasis added.)
    - Table 3-3 on "Summary of evidence for a likely to be causal relationship between long-term ozone exposure and respiratory effects" cites the study of Moore et al. (2008) ("Ambient ozone concentrations cause increased hospitalizations for asthma in children: An 18-year study in Southern California") as providing "key evidence" for the ISA's causal determination that there is "a likely to be causal relationship between long-term ozone exposure and respiratory effects." Specifically, Moore et al. is cited as providing "***Consistent evidence of an association between long-term ozone concentrations and hospital admissions***

and ED visits for asthma." Yet, follow-up work by Moore et al. (2013) noted methodological limitations of the 2008 paper (especially, that its results may have resulted from incorrect untested modeling assumptions, rather than from information in the data) and provided and applied an improved methodology ("CMRIER" or "causal models for realistic individualized exposure rules"). A key result was that the previous significant effect of ozone was no longer found. (Moore et al. (2013) state that "The results from the original HRMSM analysis based on the continuous ozone variable estimated with the G-computation method resulted in an estimate of an increase of 1.44e-06 in the proportion of asthma-related hospital discharges for a one-unit increase in ozone. [This is the 2008 study cited in Table 3-3 of the ISA.] ***Unlike results from the HRMSM analysis with the continuous ozone variable, the CMRIER results are not significant.*** Note that the HRMSM analysis was based on G-computation estimation which ***artificially relies on untestable parametric modeling assumptions*** to estimate HRMSM parameters when the ETA assumption is violated. Thus, in this ozone study [the 2008 study cited by the ISA], ***significant results from the G-computation analysis may be a consequence of the approach taken and not solely based on the information in the data***." (Moore KL, Neugebauer R, van der Laan MJ, Tager IB. Causal inference in epidemiological studies with strong confounding. Stat Med. 2012 Jun 15;31(13):1380-404. doi: 10.1002/sim.4469.) This more recent paper is not mentioned in the ISA. The ISA cites the 2008 results as "key evidence" without noting that the authors subsequently revised them in the 2013 paper.

- Table 3-3 cites a study by Tétreault et al. as providing "Key Evidence" of "Cohort studies demonstrating an *association* with asthma development in children." The ISA then interprets this, without any detailed explanation, as "Evidence for a *likely to be causal* relationship between long-term ozone exposure and respiratory effects."(Emphases added.) Yet, in discussing potential confounding, Tétreault et al. state that "We present two confounder models in the results. The first was adjusted for sex and deprivation, whereas the second was adjusted for the same variables as well as the year of birth." The article does not mention temperature or other weather variables. (For background on the importance of confounding by temperature, see e.g., Chen et al. (2018), "Does temperature-confounding control influence the modifying effect of air temperature in ozone-mortality associations?" This article concludes that using a categorical variable (e.g., a season indicator) to control for temperature yields highly significant ozone effects

at high temperatures, but also significant residual confounding; and that adjusting for (nonlinear) effects of temperatures "substantially reduced ozone effects at high temperatures and residual confounding.") Tétreault et al. also note their "**lack of information on risk factors at the individual level** (e.g. socioeconomic status and smoking). We attempted to control for these factors with adjustments of our models using ecological deprivation variables, which are imperfect and **may result in residual confounding**." (Emphasis added.) Tétreault et al. further caution that "First, **individual exposure was modeled and not measured** through the follow-up, so the quality of the associations depends on the quality of the exposure models. All associations reported in this study were estimated according to the exposure at the centroid of the residential postal code. This assumes that children would stay at home all day. Because a large proportion of a child's day can be spent outside the home (e.g., at school), where exposure to air pollutants might differ, **misclassification bias may have been introduced in our study**. Additionally, summer average O3 levels were used to estimate annual averages. Because summer O3 levels are higher than winter levels (Environment Canada 1999) in Canada, **we may have overestimated annual average levels**. Furthermore, although postal codes circumscribe a relatively small area in urban regions, postal codes may include much larger areas in rural regions. This difference in postal code size could lead to a degree of **higher imprecision in exposure estimation** in regions of the province that are less densely populated." (Emphasis added.) The ISA does not emphasize that the exposure concentrations that it reports (e.g., "32.1 ppb mean summer ozone concentration, based on 8-h midday avg" in Table 3-3) are in fact "modeled and not measured" values and does not adjust (e.g., using appropriate errors-in-variables methods) for potential biases due to such errors. It interprets the reported association as "key evidence of a likely causal relationship" without mentioning alternative interpretations such as that it might reflect omitted confounders (e.g., temperature), residual confounding, or misclassification bias. Page 3-193 of the ISA states that "Sensitivity analyses with alternate specifications for potential confounding inform the stability of findings and aid in judgments of the strength of inference from results." But it is not clear how or whether the ISA considered the results of such sensitivity analyses for the individual studies it relies on for its conclusions (e.g., in interpreting the Tétreault et al. study as "Key Evidence" of a "likely to be causal" relationship) or how sensitive the resulting causal determinations are to incompletely controlled confounding.

These examples are intended to be illustrative rather than comprehensive. They illustrate a larger issue: such coverage suggests that the draft ISA cannot serve as a trustworthy source for an accurate, unbiased, comprehensive critical summary and synthesis of the relevant scientific literature. As noted by multiple external experts and in public comments, the draft ISA appears to be biased toward defending EPA's methods and conclusions rather than providing a neutral, accurate review and summary and critical analysis and synthesis of available scientific studies. Its appearance of cherry-picking and bias in reporting results from the scientific literature undermines the effectiveness, trustworthiness, and usefulness of the draft ISA.

- *Policy-relevant science is not addressed.* The draft ISA does not usefully summarize, or critically evaluate, available scientific information on manipulative causation (i.e., on whether or to what extent reducing ozone reduces public health risks). Yet, this is the main topic needed to inform policy decisions about the public health consequences of alternative possible policy choices. For example, the external experts were directly asked "*Can valid determinations of manipulative or interventional causation – that is, how and whether changing exposure would change health risks – be made based on observed associations of the types analyzed in the ISA?*" Most who answered said no; none said yes (see responses in Appendix A). For example, Dr. North stated that "I think this is a clear **NO**. CASAC should be seeking to evaluate manipulative or interventional causation, that is, determining how many people might be added or subtracted from having their health protected with an adequate margin of safety by a change in the primary NAAQS standard." Unless this omission is fixed, the PA lack a scientific foundation in the ISA for predicting effects on public health of alternative policies.
- *Uncertain relevance of facts addressed.* The draft ISA identifies several associations between ozone and physiological changes in controlled human experiments and epidemiological data, but it does not adequately address the extent to which these associations predict adverse effects on public health. Mr. Jansen frames the issue as follows: "In addition to the issue of beneficial effects, there is the issue of recovery or reversibility. …I did not see how it affected weighting nor causality classification. In other words, if a metric was responsive but recovered, how is that evidence weighted and used in terms of causality classification?" Similarly, Dr. North states that "It seems to me an important issue whether observed mild, apparently reversible effects such as changes in FEV1 (forced expiratory volume in one second) seen in healthy young exercising subjects imply a potential for adverse health effects in the general population. What are the adverse health effects, and how well do FEV1 changes predict them? What is the C-R relationship, not just for FEV1 changes, but for adverse health impacts that are persistent

and perhaps cumulative over time, such as scarring of lung tissue so that lung function is permanently lost?" Health effects aside, the relevance of other information presented in the draft ISA is also often unclear. As stated by Dr. Parrish: "In reading through Section 1.3 a great deal of scientific information is summarized, but there is little or no discussion of the relevance of this science to the NAAQS or the ozone design values upon which the NAAQS is based." The final ISA should directly address the questions of the relevance of reversible effects, and of other information presented, to predicting public health responses to changes in ozone.

As mentioned above, the following additions to the draft ISA and Executive Summary are recommended to improve the communication of key results, and also the policy relevance, scientific validity, and methodological integrity of the content being communicated:

1. *Summarize available empirical evidence on how changes in public health effects depend on changes in ozone levels.*
2. *Present summary results from a systematic review and critical evaluation and synthesis of relevant studies*, including negative ones that have been omitted from the draft ISA.
3. *Provide detailed discussion of possible confounding*, and how it was or was not addressed for each study used to support causal conclusions.
4. *Present results of systematic evaluations of study quality*, using consistently applied criteria, showing how each key study included performs on each specific quality criterion relevant for drawing valid causal conclusions.
5. *Discuss causal biological mechanisms of inflammation-related health effects* preventable by reducing current ozone levels.
6. *Present comprehensive, quantitative uncertainty and sensitivity analyses* showing how the ISA's conclusions change for variations in selection and weighting of studies, modeling choices and assumptions, interpretations of undefined and vague terms, and subjective judgments on which the conclusions depend.

## Additional Comments on Executive Summary

p. 1 "*This Integrated Science Assessment (ISA) is a* ***comprehensive evaluation and synthesis of the policy-relevant*** *science aimed at characterizing the health and welfare effects caused by ozone.*"

*Comments:*

- The draft ISA is not a comprehensive evaluation and synthesis. Multiple public comments and external expert consultant comments have pointed out that the draft ISA omits many relevant studies and topics; it appears to multiple reviewers to exhibit a selection bias favoring including positive studies while excluding negative ones, and does not address causal C-R functions for changes in public health risks caused by changes in ozone levels.
- The draft ISA does not include policy-relevant science, i.e., studies and empirical testing and validation of predictive generalizations that would allow changes in public health risks caused by alternative changes in NAAQA to be assessed, and uncertainties about them to be characterized.

p. 1 "***The ISA identifies and critically evaluates the most policy-relevant scientific literature across scientific disciplines****, including epidemiology, controlled human exposure studies, animal toxicology, atmospheric science, exposure science, vegetation studies, agricultural science, ecology, and climate-related science.* ***Key scientific conclusions (i.e., causality determinations; Section ES.4) are presented*** *and explained. They provide the scientific basis for developing risk and exposure analyses, policy evaluations, and policy decisions for the review. This ISA draws conclusions about the causal nature of the relationships between ozone exposure and health and welfare effects by integrating information across scientific disciplines and building off the evidence base evaluated in previous reviews.* ***The ISA thus provides the policy-relevant scientific information*** *that supports the review of the NAAQS.*"

*Comments:*

- *The highly relevant disciplines of health risk analysis, decision science, causal analysis, data science, and mathematical and simulation modeling are not adequately represented or used in the draft ISA.* For example, validation of health effect models is not discussed.
- *The causality determinations are not "key scientific conclusions."* As discussed in more detail in several places in this document and in multiple public comments and external expert comments, the causal determinations appear to be only ambiguous expressions of subjective feelings and judgments, with no clear empirical implications or well-defined empirical and operational meanings. They do not provide a valid objective scientific basis for developing risk and exposure analyses, policy evaluations, or policy decisions. Referring them repeatedly as "scientific" does not make them scientific. To be genuinely

scientific, conclusions should rest on reproducible results of tests of unambiguously stated predictions against data. Neither this draft ISA nor the previous ISA for ozone presents conclusions that are "scientific" in this traditional sense of being tested against data (and modified as needed based on the results). Rather, these ISAs have appropriated the term "scientific" for opinions and judgments without applying the methods of empirical testing and repeated validation or modification of conclusions (e.g., predictive generalizations and models) that are necessary to make conclusions genuinely scientific and reliable. This should be fixed in the final ISA.

## Non-Member Consultant Responses to Questions from Dr. Cox

Following are questions (with some paraphrased explanations) sent to our external expert consultants, together with some responses received from them (not all consultants responded to all questions) and with some additional comments inserted. Emphases (bold) have been added throughout. Proposed one-word answers summarizing many of the responses (e.g., yes, no, or unclear) are provided, but the verbatim responses are frequently longer and more nuanced.

*Is the scientific information provided by the ISA clear?* (*Does it provide testable predictions, expressed using unambiguous terms with stated operational definitions? Do the conclusions and generalizations presented follow from data and evidence using transparent reasoning and derivations? Are the implications of relevant scientific knowledge for effects of actions made clear?)*

*Answer:* No.

*Explanation:* Criteria for selecting and weighting studies, and the bases for drawing conclusions from them, are not clear. It is unclear how, if at all, the conclusions would change if consistent criteria were systematically applied for selecting, summarizing, and synthesizing studies. Causal determinations appear to be ambiguous and subjective: different people might make them in different ways, with no objective basis for determining from data which (if any) is right. What they mean or imply empirically is unclear. The treatment of wildfire contributions to ozone exposure, and implications for NAAQS, are unclear.

*Responses from external consultants:*

- Mr. Jansen: "I have always been frustrated by what I perceive as a **lack of clear criteria and transparency** in the descriptions of what leads to a particular causality classification. I can read several descriptions of evidence and am **unable to identify what makes one "causal" and another "likely" or "suggestive"** and as a result have a difficult time deciding whether I agree or not." "**I have always been concerned about the term causal vs. the term contribute.** We know death certificates are problematic with primary, secondary, and even tertiary causes listed. And I know from personal experience, **they are not necessarily correct**. **How does this factor into the analysis and messaging?**"
- Dr. North: "**Is the ISA *clear*? I do not find it so to me.**" **"It is not clear in the text whether these 'design values' include wildfire episodes**, or whether wildfire episodes with even higher ozone observations were removed as "natural disaster" exceptions." "The classification in **the causal categories are judgment calls. Different people might make them differently** based on the same supporting evidence." "I… have serious reservations about basing a risk assessment for health effects on the causal determinations from the

framework EPA has used in this draft ISA. **I have difficulty interpreting what 'likely to be causal' means** in connection with possible confounding."

- Dr. Parrish: "In reading through Section 1.3 **a great deal of scientific information is summarized, but there is little or no discussion of the relevance of this science** to the NAAQS or the ozone design values upon which the NAAQS is based. **For example, wildfires** (here included in the broader category of landscape fires) and stratosphere-troposphere exchange are discussed. These are two natural sources of ozone that are specifically addressed in EPA's exceptional events rule. A crucial issue is the extent to which these sources can affect the ozone design value and perhaps cause an exceedance of the NAAQS that would be eligible for addressing under that exceptional events rule. How often do either of these natural sources cause exceedances? It is difficult to evaluate the significance of a particular scientific issue without the context of how that issue might affect the NAAQS that is being reviewed."
- Dr. Rhomberg: **"The treatment of "causation" in the ISA tends to be fuzzy** – fuzzier in some places more than in others – about the distinctions noted above. If patterns of association are plausibly explained by underlying causation, this is taken as sufficient evidence for such causation (when one should actually be comparing the hypothesized causative actions against other competing explanations for the patterns), and if such causation is inferred, it is taken to be universal, applying to other settings on the usually poorly stated (much less justified) presumption that the causation is universal and largely independent of other circumstances. The **measures of association are too easily taken as measures of the magnitude of cause, and these magnitudes applied to other settings without due consideration of how they may be contingent on differing circumstances**."
- Dr. Sax: "In my opinion, **the ISA could be clearer in many aspects**… Given the same information, I believe, that **different people (or different groups) could come to very different conclusions** and causal determinations based on current classification descriptions." "In addition, **the ISA causality framework would also benefit from some clarity regarding what each causal classification represents**. Specifically, is the classification meant to identify the strength of the evidence for a given level of exposure (e.g., below the current NAAQS) or is it simply to identify whether there is evidence enough of expected harm at any level of exposure?"

*Is the scientific information provided by the ISA sound?* Is it logically valid, correctly stated with appropriate caveats and based on the total body of relevant evidence?

*Answer:* No.

*Explanation:* The Draft ISA misinterprets evidence of association as evidence of causation without presenting data and analyses that rule out other interpretations (such as confounding, temporal trends, modeling choices, study selection biases, data aggregation biases, measurement error biases, etc.) It does not show that reducing ozone has caused or would cause any change in public health risks.

*Responses from external consultants:*

- Mr. Jansen: "I am of the school that a) **associations are not and cannot be causal** and b) quality human and animal experimental studies at relevant exposures need to be weighted over suggestive epidemiological (associational) studies to establish causality." (Tony Cox comment: This school of thought is discussed further by Pearl 2009, 2010. It holds that causal and associational concepts "do not mix" because causation involves determining how changing one variable changes others, while association only describes whether and how observed levels of variables co-vary.)
- Dr. Lipfert: "**The ultimate test of causality is whether health has actually improved** since the late 1970s in response to peak O3 levels reduced by a factor of 5 in conjunction with coincident trends in spatial patterns of reduced smoking and improved medical care. **A search of PubMed found no support for such improvement.**"
- Dr. North: "**Is the ISA *sound*? Not in my judgment,** because there is little exploration of confounding and **too much reliance on strength of association** in the Bradford Hill criteria. I thought Appendix 2 [put] **insufficient emphasis on the differences between human exposure indoors, where most people spend most of their time, and ambient levels outdoors**. I found **weakness with the discussion of inflammation** in clinical studies at near-ambient exposure levels in Appendix 3, and I am **dissatisfied generally with the discussions of causality and confounding**."
- Dr. Parrish: "Section 1.3.1.3.3 **Landscape Fires deserves improvement.** It provides a synopsis of literature results, but there is no synthesis of the current state of knowledge. …**If wildfires have relatively large, episodic effects, why are these effects not seen in the ozone design values** in these rural northern U.S. states where wildfire impacts are expected to be most obvious?"
- Dr. Sax: "In general, **it is difficult to follow the rationale as presented in the ISA with regards to causal conclusion determinations** and **there are appear to be inconsistencies** in how the evidence is deemed sufficient to select one classification vs. another, such as in the cardiovascular vs. the metabolic outcomes. This highlights the need to have clear protocols and definitions regarding how these conclusions are derived."

*Are the key scientific conclusions provided by the ISA scientific?* Are they falsifiable by observations, tested against data, independently verifiable and verified, and logically sound?

*Answer:* No.

*Explanation:* The Draft ISA does not show the results of testing its conclusions and predictions against data not used in deriving them, i.e., predictive validity has not been established. As noted above, its key conclusions (causal determinations) appear to be subjective judgments expressed using ambiguous language: other people might make different judgments based on the same data, and there is no clear way to determine which judgment (if any) is right. Uncertainty about these vague, subjective judgments is not characterized.

*Responses from external consultants:*

- Dr. North: "Is the ISA *scientific*? The classification in **the causal categories are judgment calls. Different people might make them differently** based on the same supporting evidence."
- Dr. Sax: "The operational definitions of the causal **classifications leave a lot of room for subjective interpretation**."

*Is the scientific information provided by the ISA policy-relevant?* Does it quantify how changing ozone NAAQS would change (probable) public health outcomes, characterizing uncertainties, interactions, modifiers, and confounders?

*Answer:* No.

*Explanation:* The Draft ISA does not characterize whether or to what extent reducing ozone would cause (or has caused) lasting effects on public health risks.

*Responses from external consultants:*

- Mr. Jansen: "In addition to the issue of beneficial effects, there is the issue of recovery or reversibility. I did a search for the terms in the ISA and found some references to them in discussing experimental studies, which is appropriate. However, I did not see how it affected weighting nor causality classification. **In other words, if a metric was responsive but recovered, how is that evidence weighted and used in terms of causality classification?**"
- Dr. Lipfert: "My overall conclusion from the above is that **a risk-free NAAQS for ozone cannot be determined at any level above background** and certainly not as high as 70 ppb. It thus follows that **society must determine a tolerable level of risk, taking into account the presence of extreme frailty. This situation was not considered in the framing of the Clean Air Act**."

- Dr. North: "It seems to me an important issue **whether observed mild, apparently reversible effects** such as changes in FEV1 (forced expiratory volume in one second) seen **in healthy young exercising subjects imply a potential for adverse health effects in the general population**. What are the adverse health effects, and how well do FEV1 changes predict them? **What is the C-R relationship, not just for FEV1 changes, but for adverse health impacts that are persistent** and perhaps cumulative over time, such as scarring of lung tissue so that lung function is permanently lost?"
- Dr. Parrish: "In reading through Section 1.3 **a great deal of scientific information is summarized, but there is little or no discussion of the relevance of this science** to the NAAQS or the ozone design values upon which the NAAQS is based. For example, wildfires (here included in the broader category of landscape fires) and stratosphere-troposphere exchange are discussed. These are two natural sources of ozone that are specifically addressed in EPA's exceptional events rule. A crucial issue is the extent to which these sources can affect the ozone design value and perhaps cause an exceedance of the NAAQS that would be eligible for addressing under that exceptional events rule. How often do either of these natural sources cause exceedances? It is difficult to evaluate the significance of a particular scientific issue without the context of how that issue might affect the NAAQS that is being reviewed."

*Can valid determinations of manipulative or interventional causation – that is, how and whether changing exposure would change health risks – be made based on observed associations of the types analyzed in the ISA?*

*Answer:* No.

*Explanation:* Associations between past levels of variables do not predict how changing one variable would change another. Such predictions require causal analysis (Pearl 2009, 2010).

*Responses from external consultants:*

- Dr. Lipfert: "**No**." "**The ultimate test of causality is whether health has actually improved since the late 1970s in response to peak O3 levels reduced by a factor of 5** in conjunction with coincident trends in spatial patterns of reduced smoking and improved medical care. **A search of PubMed found no support for such improvement**."
- Dr. North: "I think this is a clear **NO**. CASAC should be seeking to evaluate manipulative or interventional causation, that is, determining how many people might be

added or subtracted from having their health protected with an adequate margin of safety by a change in the primary NAAQS standard."

- Dr. Rhomberg: "**It is not possible to make an undisputable, totally sufficient conclusion of interventional causation,** because of the INUS nature of the possible causative processes." (Note: Dr. Rhomberg also explains that "The way these issues affect the 'causality' problem is that the complexities of context and potential interactions are what lead to the difficulty in inferring between the different aspects of causality – ability to affect, responsibility for observations, ability to generalize effect to other settings, ability to measure the impact of an effect, and ability to generalize that magnitude to other settings. They also affect the certainty with which causal determinations can be made, constituting the things that need to be thought through in characterizing uncertainty in any causation assertions." Dr. Rhomberg suggests that the NAS might provide useful advice on how methodological issues for causal determination framework can best be addressed.)
- Dr. Sax: "The short answer to this question I think is **no**."
- Dr. Thomas: "However, **the vast bulk of air pollution studies have not been designed or analyzed for the purpose of assessing manipulative or interventional causation**." *Note:*Dr. Thomas adds that "Nevertheless, the consistency of the findings from numerous observational studies, the concordance with short term human experimental studies (e.g., chamber or panel studies), and animal experiments, along with other lines of evidence supporting biological plausibility, as outlined in the preface to the ISA, allows a causal interpretation in terms of the likely effect of air pollution on the various health endpoints, if not a quantitative estimate of the predicted magnitude of the effect of a hypothetical intervention." However, the claimed consistency of associations is affected by modeling choices, study selection choices, averaging of disparate results, consistency in omission or incomplete control of confounders, and other factors that do not warrant a causal interpretation. Moreover, "a quantitative estimate of the predicted magnitude of the effect of a hypothetical intervention" is then in fact based on these associations in the PA.

*Is this actually a "formal causal framework*"?

- Dr. Lipfert: "**No**, it's a list of subjective rationalizations based on studies selected from the literature according to unspecified procedures."
- Dr. North: "**No.** I agree that the terms are not clearly and unambiguously defined."
- Dr. Rhomberg: "**A formal causal framework" implies that a path to certain determination is available, and I do not think one is.**"

- Dr. Sax: "I am unsure of the answer/ or **answer is not clear**. I generally agree that the current **definitions for each classification leave a lot of room for subjective judgement** and in general it is not always clear (to me) how EPA weighs the evidence and comes to its final causal conclusions."
- Dr. Thomas: **"The approach used in the ISA does not exploit the emerging framework of "causal inference"** that constitutes one type of "formal causal framework." However, the "weight of evidence" machinery (Committee_to_Review_the_IRIS_Process 2014) used here is certainly a well-established and appropriate formal framework for reaching causal judgements combining evidence across scientific disciplines. The machinery of statistical causal inference is not capable of or intended to synthesize evidence across multiple studies from multiple scientific disciplines. … [For] a determination of a "causal relationship" the following is required: "… Generally, the determination is based on multiple studies conducted by multiple research groups, and evidence that is considered sufficient to infer a causal relationship is usually obtained from the joint consideration of many lines of evidence that reinforce each other." (Tony Cox notes: The WoE framework is not a formal or operational framework, insofar as it specifies no empirical operations or procedure for verifying or refuting a causal determination classification. Rather, the WoE franework offers a circular definition of causality: that a "causal" determination is generally based on "evidence that is considered sufficient to infer a causal relationship." This definition lacks verifiable empirical content. The WoE framework has not been validated for ozone. No studies showing that the WoE framework is "appropriate" (e.g., performing better than, or as well as, random guessing) have been published, as far as I have been able to determine by asking Chris Frey, Lianne Sheppard, and EPA. The causal inference framework that Dr. Thomas and others describe as "emerging" actually emerged approximately 100 years ago. It includes algorithms that are capable of, and intended to, synthesize evidence across multiple studies drawing on multiple scientific disciplines and both experimental and observational data. The misconception that formal causal analysis applies only to single data sets can be corrected by googling "integrative causal analysis" to obtain points of entry to the substantial technical on how to apply formal causal modeling and analysis algorithms to integration and synthesis of diverse data sets from multiple studies or fields.)

*Does the ISA's causal determination framework clearly distinguish between necessary and sufficient causation?*

*Answer:* No.

*Responses from external consultants:*

- Mr. Jansen: "**[T]he answer is no** simply because the framework does not accommodate those terms."
- Dr. Lipfert: "**No**."
- Dr. North: "**No.** Other factors should be considered, and positive association is not the same as necessary or sufficient. One needs to think in terms of partial causation through examination of multiple factors."
- Dr. Rhomberg: "**No,** as noted above, the naïve notion of causation sort of implies both necessity and sufficiency…"
- Dr. Sax: "The short answer to this is **no**."

*Does a determination that exposure has a "causal relationship" with a health effect in a population imply that reducing exposure would reduce risk of the health effect in the population, other factors being held fixed? In other words, does a "causal relationship" determination imply a manipulative causal relationship?*

*Answer:* No.

*Explanation:* Causal determinations in the Draft ISA are based primarily on associations and biological considerations. They do not address manipulative or interventional causation.

*Responses from external consultants:*

- Dr. Lipfert: "**Not in my opinion**."
- Dr. North: "**No.** I do not find that "causal" as used in the ISA implies "a manipulative causal relationship."
- Dr. Sax: "I think again the answer is **no**. … The observed associations do not imply causation…"
- Dr. Thomas: "That depends upon the context in which the term "causal relationship" is used. In the statistical literature on causal inference, yes, the goal is to estimate the effect of an intervention on differences in the expected outcome within an individual under different hypothetical scenarios. In the epidemiological literature -- and **as used in the ISA -- it refers to the existence of a mechanism under which exposure is a contributing factor**, which may imply that a change in exposure would be expected to

change the outcome, but that is not the primary sense in which the term is used." (Tony Cox comments: "Existence of a mechanism under which exposure is a contributing factor" has no clear implications for whether or how much reducing exposure would reduce risk. Hence, it lacks clear policy relevance. Policy relevance requires addressing manipulative causation, i.e., how would a change in policy change affect public health risks? Insofar as the ISA does not address or answer this question, it does not provide the most important scientific information needed to inform policy decisions.)

*Can causal determinations be incorrect*?

*Answer:* Yes.

*Responses from external consultants:*

- Mr. Jansen: "I have always been concerned about the term causal vs. the term contribute. We know death certificates are problematic with primary, secondary, and even tertiary causes listed. And I know from personal experience, **they are not necessarily correct**. **How does this factor into the analysis and messaging?**"
- Dr. North: **Yes**.
- Dr. Rhomberg: "**Yes, it can be incorrect.** It cannot be surely correct, so the question is how usefully to characterize the uncertainty and also how to act in the face of such uncertainty…"
- Dr. Sax: "I think that the causal classifications can be interpreted differently based on the current definitions and how EPA presents and interprets the data. That is, **they are open to interpretation**, particularly for determinations of a "causal relationship.""
- Dr. Thomas: "**Yes, of course, any human judgment could be incorrect.** That is true of those reached by a large body of experts as well but is much less likely!" (Tony Cox comment: Studies of groupthink, conformation bias, and group judgment have identified many factors that can make large bodies of experts even more likely than individuals to reach incorrect judgments. However, both individual and collective judgments about causality are notoriously error-prone and unreliable, making it preferable to rely on formal analytic procedures (e.g., conditional independence tests) applied to relevant data to draw causal conclusions.

I*s it clear how uncertainty about which category is correct should be (or has been) resolved* in assigning a final causal determination category, as in Table ES-1 p. ES-5) of the ISA?

*Answer:* No.

*Responses from external consultants:*

- Mr. Jansen: "I am a proponent of a quantitative uncertainty analysis being performed. And I would argue that **the ISA should include a section <u>in each chapter</u>on the literature that evaluates the uncertainty in the various components that will make up the risk assessment.** One could look to the outline provided in Dr. Lange's question number 4 (questions to consultants on the PM PA) for the beginnings of an outline, adapted of course for ozone. I do believe substantial data and hopefully studies exist to derive estimates for many of the items in her Table 1 and EPA should get on with performing that work. They have been advised to do so in the past. The approaches recommended by Dr. Anne Smith (see below taken from PM PA comments of Dr. North) should also be tried and information embedded in the various studies needed for such analysis should be summarized."
- Dr. Lipfert: "**No**."
- Dr. North: **No.**
- Dr. Rhomberg: "I have argued that **this could be improved,** though it is hard to spell out a remedy…" (Tony Cox comment: One remedy might be to characterize the probability that each determination category applied to the available evidence. Assigning a single category when there is uncertainty about which is correct does not fully characterize relevant uncertainty for policy makers.)
- Dr. Sax: "**No, I don't think it is clear.** As noted above, EPA should provide more clarity and perhaps caveat the classifications based on the amount of uncertainty in the underlying evidence."
- Dr. Thomas: "While the process for deciding upon which category is appropriate is clearly described, **I do not see much if any discussion about [how] any disagreements about the choice of category were resolved,** except in terms of justification for changes in the categorization since the 2013 ISA (Tables IS-4 and IS-5)." (Tony Cox comment: My question was about uncertainty rather than about disagreements. Unanimous agreement on choice of category need not imply that there is no uncertainty; indeed, it might reflect a desire to conform when uncertainty is large enough to prevent easily defended disagreements.)

*Is it clear how observations could be used to test and falsify a given causal determination if it is not correct?*

*Answer:* No.

*Responses from external consultants:*

- Dr. Lipfert: "Laboratory experiments may determine whether a given type of response *can* happen. Epidemiology is required to determine under what circumstances it *does* happen."
- Dr. North: **No.**
- Dr. Rhomberg: "There are methods to try to uncover contingencies on other factors, by measuring them and using directed graph analysis."
- Dr. Sax: "**No,** I don't think that as the framework is laid out by EPA, it could be applied consistently and that the same causal determinations would be necessarily developed by different groups. As noted above, this is exemplified by the change in classification for two key health outcomes (CV and mortality) in the current ISA."
- Dr. Thomas: "While it is always possible that new data will emerge that leads one to question a previous determination, such speculation would be beyond the scope of the ISA." (Tony Cox comment: My question was whether causal determinations are falsifiable by data (including existing data), and if so, how. For example, if particulate matter in a carefully performed accountability study were permanently reduced by 30 or more micrograms per cubic meter by a ban on coal burning, and if this reduction had no observed effect on mortality rates on time scales of years to decades in multiple relevant comparison groups, would this outcome falsify a causal determination of "causal"? Why or why not?)

*Is the correctness of each causal determination in table ES-1 formally and transparently evaluated in the ISA? In other words, have formal rules for determining the correctness of the causal determinations in Table ES-1 (p. ES-5) from the data and evidence presented been explicitly stated, applied systematically, and the results documented? (If so, where?)*

*Answer:* No.

*Responses from external consultants:*

- Dr. Lipfert: "**No.** I find Table ES-1 useless and completely subjective."
- Dr. North: "**No.** I do not believe the determination represented by the entries in Table ES-1 can be defended as correct. I think the Administrator, CASAC, and we who are advising CASAC could argue *ad nausium* about causality, **and we would be better off trying to address how to estimate the extent of health response** and the severity of the health response for exposures in the 60-70 ppb range."

- Dr. Rhomberg: "I have argued **that a more rigorous conception of causation is indeed needed**. Bradford Hill considerations and other such approaches are at best guides for fuller thinking through, they are not tests to be passed."
- Dr. Sax: "**No,** see my comments regarding the NAAQS framework above, as well as Goodman et al. (2013) reference (above)."

*Does a determination that an exposure-response (or concentration-response (C-R)) relationship is a "causal relationship" imply that it is <u>entirely</u> causal*, with no contribution from incompletely controlled confounding, modeling errors and biases, or other non-causal sources?

*Answer:* No.

*Responses from external consultants:*

- Dr. North: "**No.**"
- Dr. Rhomberg: "**No."**
- Dr. Sax: "This is an interesting question, and I think that the way that EPA uses the C-R functions in the risk assessment, assumes that there is a causal relationship between the air pollutant (in this case ozone) and the various health effects, under the conditions specified in the underlying epidemiology model and then EPA often applies this to other populations. However, **the epidemiological studies are flawed in that they do not account for uncontrolled or incompletely controlled confounders**, **and have errors and biases** that are not always discussed or caveated. Therefore, it is more likely that there does exist some lower bound of the attributable fraction and it is unclear how much of it is necessary for consideration of a true causal effect."
- Dr. Thomas: "**No, but that depends upon what is meant** by "entirely causal." Epidemiologists have long recognized a "complex web of causation (MacMahon and Pugh 1970) meaning that no single factor is ever both necessary and sufficient to cause disease. A "causal relationship" is generally held to mean that a risk factor is a real component of one of the "sufficient component causes" of disease (Rothman 1976)."

*Does a determination that a C-R relationship is a "causal relationship" imply <u>100% certainty</u> that it is causal?*

*Answer:* Unclear.

*Responses from external consultants:*

- Dr. North: "**No.** Let's use probabilities instead of seeking yes-no answers."

- Dr. Rhomberg: "**No**"
- Dr. Sax: "I think as EPA uses it, **yes**."
- Dr. Thomas: "Obviously **that depends** upon the confidence with which that judgment has been reached." (Tony Cox note: The confidences with which causal determination judgments in the ISA have been reached are not specified.)

*Does a determination that a C-R relationship is a "causal relationship" imply that it is causal for <u>every member</u> of a population,* or might it be deemed "causal" if it is causal for a sensitive subpopulation only?

*Answer:* No.

*Responses from external consultants:*

- Dr. North: "**No.** Sensitive subpopulations must be considered."
- Dr. Rhomberg: "**No**"
- Dr. Sax: "EPA currently uses C-R functions in its risk assessment based on selected epidemiological studies. The epidemiological studies are based on different population groups (sometimes children, sometimes the elderly, sometimes all ages etc.), therefore I think **it depends on the epidemiological stud**y and the population group that the study includes."
- Dr. Thomas: "**Not necessarily**."

*Are the five categories mutually exclusive?*

*Answer:* Unclear.

*Responses from external consultants:*

- Dr. North: "I must respond that the lines between the categories are **unclear to me**. I cannot attest to mutually exclusivity. Again, I think the level of ozone exposure is critical. CASAC should evaluate the significance to health (and welfare) of ozone exposures in the 60-100 ppb range, in the context of other factors. Causality for adverse health effects seems well established at higher levels in the occupational health literature, for example, the Ozone Material Safety Data Sheets. See https://www.cdc.gov/niosh/idlh/10028156html; www.amsbio.com/images/featureareas/ozilla/Ozilla-MSDS.pdf."
- Dr. Rhomberg: "**Unclear**, partly because the meaning of each category could be tweaked. What is clear is that the categories do not seem well suited for making the distinctions

between existence of some causal property, the degree of independence of that property from other circumstances, and the possibilities and limits of generalization."

- Dr. Sax: **"I don't think they are necessarily mutually exclusive**. Again, as defined I think the scientific evidence could be evaluated differently by different people or groups such that they would arrive at different conclusions (or different categories) given the same evidence. See for example our assessment of the cardiovascular effects of ozone and EPA's conclusions in the prior ozone ISA. Prueitt, RL; Lynch, HN; Zu, K; Sax, SN; Venditti, FJ; Goodman, JE. 2014. "Weight-of-evidence Evaluation of Long-term Ozone Exposure and Cardiovascular Effects." Crit. Rev. Toxicol. 44(9):791-822. Goodman, JE; Prueitt, RL; Sax, SN; Lynch, HN; Zu, K; Lemay, JC; King, JM; Venditti, FJ. 2014. "Weight-of-evidence Evaluation of Short-term Ozone Exposure and Cardiovascular Effects." Crit. Rev. Toxicol. 44(9):725-790."
- Dr. Thomas: "**Yes.**"

*Are the five categories collectively exhaustive?*

*Answer:* Unclear.

*Responses from external consultants:*

- Dr. North: "I am not sure I care about "collectively exhaustive." I do not think the framework is useful."
- Dr. Rhomberg: **Unclear**
- Dr. Sax: "**Not necessarily**, as there is no counter classification to each of the levels"
- Dr. Thomas: "**Yes.**"

*Can a body of evidence be categorized as "likely to be causal" if the probability of causality based on the evidence is less than 50%?*

*Answer:* Unclear.

*Responses from external consultants:*

- Dr. North: "I prefer probability statements and partial causation with multiple factors as in Dr. Cox's example. I do not like implied "bright lines," such as greater than 50% means likely. Such conventions need to be agreed to among the users. I do not believe that is the case here."
- Dr. Sax: "Based on EPA's definitions, it is **hard to say** what percent probability of causality needs to be for the evidence to be considered "likely to be causal.""

- Dr. Thomas: "Causal inference methods aim to estimate the "Average Causal Effect", not the probability of causality.The "Probability of Causation" (PC) is an estimate of the probability that a specific individual's disease was caused by some aspect of his exposure history, essentially an individualized version of the epidemiologic concept Population Attributable Risk Fraction." (Tony Cox comment: My question was about probability of causality for a body of evidence in a population, and what level justifies a label of "likely to be causal," not about probability of causation for an individual.)

*Is it clear that the ISA's study selection process has successfully provided a comprehensive, trustworthy, and unbiased selection of the best available science on ozone and health effects?* Is it clear why results from Moore et al. (2008) are included and cited as "key evidence" but contrary results from Moore et al. (2013) are excluded? More generally, *is it clear that study inclusion and exclusion criteria were applied systematically and neutrally to identify and select the best and most up-to-date studies to inform the ISA's conclusions?*

*Answer:* No.

*Responses from external consultants:*

- Mr. Jansen: "**Clear criteria are needed for study inclusion/exclusion**, study quality, and causality classification. It is also **not always clear which evidence is being given more weight** than other evidence." "[In the past], when significant negative and null studies were included there was a tendency for the narrative to critique (dismiss?) them and simply accept the significant positive studies. **This suggests a bias in the review of the evidence.** Identified missing studies also ran the risk of post-rationalization. EPA needs to work harder to critique all studies, weight them appropriately, and avoid post-rationalization of additional studies identified in the review process. Unfortunately, this bias is continuing."
- Dr. Lipfert: "As evidence of **subjective selection of epidemiology studies**, I note that only one of my studies is included (see Table 1 above), but the authors saw fit to mention a study of O3 exposure and erectile disfunction. …Studies with wide confidence intervals were included."
- Dr. North: "I and others have found **other relevant studies not included**. There is a deficit of published papers on interpretation of the available studies, especially on uncertainty, variability and severity of health effects. Discussion of inflammation is an example. Non-experts have difficulty understanding the importance of the range of biomarkers for assessing the degree to which these biomarkers indicate a public health impact deserving of protection under the language of the Clean Air Act. I did not find any

references to the journal *Risk Analysis*, for which I am an area editor, and which publishes many papers on air pollution health risks. I expect other journals may be neglected as well."

- Dr. Parrish: "Another significant shortcoming in Section 1.3 is that the **uncertainty of the ozone precursor emissions estimates should be clearly discussed and defined** to the extent possible."
- Dr. Sax: "In my opinion, **the ISA could be clearer in many aspects, including how studies are selected for inclusion (or exclusion)**, and **how the evidence is weighed** and specifically used to reach causal conclusions. …**The ISA could benefit from being clearer on how the literature is selected (i.e., more specifics on search criteria used, screening on literature, literature included and excluded, and reasons for inclusion and exclusion).** This would provide greater transparency related to this process."
- Dr. Thomas: "I agree that the **Moore (2013) should be discussed**."

*Are there other studies that are omitted from the ISA that should be included?*

*Answer:* Yes.

*Responses from external consultants:*

- Dr. Lipfert: "I summarized ozone mortality risk estimates for both long-and short-term exposures from my publications and from the literature in Table 1 (note that **my publications have not been included in EPA ISAs**. In general, there is a lot of heterogeneity among these estimates, even within the same cohort."
- Dr. North: "I am concerned that **EPA's selection process is leaving out studies with negative findings for ozone**. This is evidence against the ISA study selection process being comprehensive, trustworthy, and unbiased." "I would like to see possible interventions made explicit at the regional level and evaluated in terms of **how much these interventions would reduce exposure, and then, what impacts these would have on health and welfare measures**, with results in probabilistic form. **Anne Smith's 2019 paper in *Risk Analysis*** demonstrates how such an evaluation might be done." "I did not find **Kai [Chen] et al. (2018) in** the references in ISA Appendix 3, and that concerns me a great deal [https://journals.lww.com/environepidem/FullText/2018/03000/Does_temperature_confounding_control_influence_the.2.aspx]."
- Dr. Parrish: "[T]here is evidence (Oikawa et al., 2015; Almaraz et al., 2018) **that NOx emissions from agricultural soils plays a major role in degraded air quality in the**

**Imperial Valley, located in an ozone no-attainment area in California. This issue deserves more discussion** since it may play a significant role in an important no-attainment areas in the country."

- Dr. Sax: "**[S]everal studies** (some of which I am an author of and references provided throughout) **are missing from the ISA. More transparency is needed on how EPA selected studies for inclusion/exclusion**." "I refer Dr. Frampton to several studies that were published by me and my colleagues, and the findings from these evaluations are more consistent with the change in the causality determination for CV disease. **None of these studies were mentioned or included in the ozone ISA**: Prueitt, RL; Lynch, HN; Zu, K; Sax, SN; Venditti, FJ; Goodman, JE. 2014. "Weight-of-evidence Evaluation of Long-term Ozone Exposure and Cardiovascular Effects." Crit. Rev. Toxicol. 44(9):791-822. Goodman, JE; Prueitt, RL; Sax, SN; Lynch, HN; Zu, K; Lemay, JC; King, JM; Venditti, FJ. 2014. "Weight-of-evidence Evaluation of Short-term Ozone Exposure and Cardiovascular Effects." Crit. Rev. Toxicol. 44(9):725-790. Goodman, JE; Prueitt, RL; Sax, SN; Pizzurro, DM; Lynch, HN; Zu, K; Venditti, FJ. 2015. "Ozone Exposure and Systemic Biomarkers: Evaluation of Evidence for Adverse Cardiovascular Health Impacts." Crit. Rev. Toxicol. 45(5):412-452. Petito Boyce, C; Goodman, JE; Sax, SN; Loftus, CT. 2015. "Providing Perspective for Interpreting Cardiovascular Mortality Risks Associated with Ozone Exposures." Reg. Tox. Pharmacol. 72(1):107-116."
- Dr. Thomas: "I do not have the comprehensive knowledge of the air pollution epidemiology literature to address whether there are key studies missing, other new evidence, or inadequately considered weaknesses."

*Are there studies included in the ISA that should be omitted* (e.g., because of uncontrolled confounding, obsolete or incorrect modeling assumptions, conclusions dependent on unverified assumptions, ecological fallacy, lack of causally relevant information, lack of design that can support valid causal inferences, or other methodological problems?)

- Dr. Sax: "Other than general study quality criteria that EPA has identified in an Annex to each of the Appendices, it is still unclear how EPA weighs studies based on study quality**.** **This is an area that EPA still needs to work on and add to its ISA evaluations**. After an assessment of study quality, which would help to identify methodological issues or biases in certain studies, EPA **could develop criteria for excluding studies** or at the very least for giving certain studies less weight in the overall conclusions and causal determinations."

*Is it clear that the process followed in selecting and summarizing scientific studies in the ISA was sufficient to assure accurate, unbiased, up-to-date, and trustworthy summaries of the relevant scientific literature to inform causal determination judgments?*

*Answer:* No.

*Explanation:* Neither the selection of studies nor the summary of studies that have been selected appears to be unbiased, comprehensive, or trustworthy. For example, I searched PUBMED for "ozone respiratory effects causal." Three of the top seven articles returned are as follow, shown here with selected conclusions (emphases added):

- Qian Z, He Q, Lin HM, Kong L, Zhou D, Liang S, Zhu Z, Liao D, Liu W, Bentley CM, Dan J, Wang B, Yang N, Xu S, Gong J, Wei H, Sun H, Qin Z; HEI Health Review Committee. Part 2. Association of daily mortality with ambient air pollution, and effect modification by extremely high temperature in Wuhan, China. Res Rep Health Eff Inst. 2010 Nov;(154):91-217. "Among the gaseous pollutants, we also observed statistically significant associations of mortality with NO, and SO2, and that the estimated effects of these two pollutants were stronger than the PM10 effects. The patterns of NO2 and SO2 associations were similar to those of PM10 in terms of sex, age, and linearity. ***O3 was not associated with mortality***."
- Cox LA Jr, Popken DA. Has reducing fine particulate matter and ozone caused reduced mortality rates in the United States? Ann Epidemiol. 2015 Mar;25(3):162-73. doi: 10.1016/j.annepidem.2014.11.006. "***There were no significant positive associations between changes in PM2.5 or O3 levels and corresponding changes in disease mortality rates*** between 2000 and 2010, nor for shorter time intervals of 1 to 3 years."
- Goodman JE, Prueitt RL, Chandalia J, Sax SN. Evaluation of adverse human lung function effects in controlled ozone exposure studies. J Appl Toxicol. 2014 May;34(5):516-24. doi: 10.1002/jat.2905. "Overall, ***these studies do not demonstrate a causal association between ozone concentrations in the range of the current National Ambient Air Quality Standard and adverse effects on lung function***."

None of these negative results is mentioned in the draft ISA.

For studies that that *are* cited in the ISA, I performed the following spot checks and found the following results.

- Page 3-91 of the ISA states that "A limited number of recent studies provide evidence of an association between long-term exposure to ozone and asthma development in children. … An overview of the evidence is provided below. A recent CHS analysis examined asthma incidence in relation to improved air quality in nine southern California communities (Garcia et al., 2019). ***Decreases in baseline ozone concentrations in three CHS cohorts, enrolled in 1993, 1996, and 2006, were associated with decreased asthma incidence.***" However, Garcia

et al. (2019) actually state that "Among children in Southern California, decreases in ambient nitrogen dioxide and PM2.5 between 1993 and 2014 were significantly associated with lower asthma incidence. ***There were no statistically significant associations for ozone*** or PM10." (Garcia E, Berhane KT, Islam T, McConnell R, Urman R, Chen Z, Gilliland FD. Association of Changes in Air Quality With Incident Asthma in Children in California, 1993-2014. JAMA. 2019 May 21;321(19):1906-1915. doi: 10.1001/jama.2019.5357. Emphasis added.)

- Table 3-3 on "Summary of evidence for a likely to be causal relationship between long-term ozone exposure and respiratory effects" cites the study of Moore et al. (2008) ("Ambient ozone concentrations cause increased hospitalizations for asthma in children: An 18-year study in Southern California") as providing "key evidence" for the ISA's causal determination that there is "a likely to be causal relationship between long-term ozone exposure and respiratory effects." Specifically, Moore et al. is cited as providing "***Consistent evidence of an association between long-term ozone concentrations and hospital admissions*** and ED visits for asthma." Yet, follow-up work by Moore et al. (2013) noted methodological limitations of the 2008 paper (especially, that its results may have resulted from incorrect untested modeling assumptions, rather than from information in the data) and provided and applied an improved methodology ("CMRIER" or "causal models for realistic individualized exposure rules"). A key result was that the previous significant effect of ozone was no longer found. (Moore et al. (2013) state that "The results from the original HRMSM analysis based on the continuous ozone variable estimated with the G-computation method resulted in an estimate of an increase of 1.44e-06 in the proportion of asthma-related hospital discharges for a one-unit increase in ozone. [This is the 2008 study cited in Table 3-3 of the ISA.] ***Unlike results from the HRMSM analysis with the continuous ozone variable, the CMRIER results are not significant.*** Note that the HRMSM analysis was based on G-computation estimation which ***artificially relies on untestable parametric modeling assumptions*** to estimate HRMSM parameters when the ETA assumption is violated. Thus, in this ozone study [the 2008 study cited by the ISA], ***significant results from the G-computation analysis may be a consequence of the approach taken and not solely based on the information in the data***." (Moore KL, Neugebauer R, van der Laan MJ, Tager IB. Causal inference in epidemiological studies with strong confounding. Stat Med. 2012 Jun 15;31(13):1380-404. doi: 10.1002/sim.4469.) This more recent paper is not mentioned in the ISA. The ISA cites the 2008 results as "key evidence" without noting that the authors subsequently revised them in the 2013 paper.

*Responses from external consultants:*

- Dr. North: "I am concerned that **EPA's selection process is leaving out studies with negative findings for ozone**. This is evidence against the ISA study selection process being comprehensive, trustworthy, and unbiased. Including Moore (2008) but not Moore (2013) is further evidence of weakness…"

*Do you find in the Executive Summary a clear explanation of the extent to which the key evidence supporting the ISA's causal determinations consists of, is sensitive to, or is derived from unverified modeling assumptions, or from modeling assumptions that more recent literature has found to be incorrect or inadequate? Have you found information in the ISA on sensitivity of causal determination conclusions to untested, uncertain, or incorrect assumptions? (If so, where? See Table Annex 6-1, cf p. 6-67 for a discussion of what should be done. Has it be done, and is it clear what the results were?)*

*Responses from external consultants:*

- Dr. North: "I am disappointed that the Executive Summary focuses on the causality determination, rather than a common sense summary of the relevant science including discussion of modeling assumptions, uncertainties and variability in ozone exposure. I did not find the Annex to Appendix 6 useful as a guide to "what should be done." It seemed like a defense of EPA practice as best practice, and I disagree, especially on the use of the Bradford Hill criteria."

*Is it clear that the individual studies cited in support of the ISA's causal determinations of "causal" or "likely to be causal" adequately controlled for potential confounding and residual confounding by variables such as income and weather variables?*

*Answer:* No.

*Responses from external consultants:*

- Dr. North: "Issues of **how to deal with confounding by temperature, extent of air conditioning in homes and workplaces, and socioeconomic status – interrelated factors that will differ by location – need to be carefully evaluated in order to get good estimates of mortality and morbidity responses. EPA in this ozone ISA seems far behind evolving 'best practices'** in how to do such analysis."
- Dr. Thomas: "Of those studies cited in support of these determinations that I am familiar with, the authors have gone to appropriate lengths to control for such confounders, to the extent possible with the available data. **Of course, residual confounding can never be excluded** from any observational epidemiology study." (Tony Cox comment: Residual

confounding by variables such as income and temperature can easily be excluded by using continuous values instead of categorizing them.)

*Is it clear that the individual studies cited in support of the ISA's causal determinations of "causal" or "likely to be causal" have adequately controlled for biases due to exposure estimation errors or exposure misclassification errors?*

*Answer:* No.

*Responses from external consultants:*

- Dr. North: "Estimating exposure 'according to the centroid of the postal code' is quite crude, and misses all the subtleties whether children are outdoors, indoors in air conditioned space, or indoors in non-air conditioned space, and the extent to which they are exercising in these environments. It also may miss exposure to materials the aggravate asthma, such as pet and cockroach dander."
- Dr. Thomas: "While the various studies differ in their methods of exposure estimation, **few to my knowledge have used formal methods of measurement error correction.** However, the bias from measurement error would generally be in the direction of reducing effect sizes and power, not introducing false positives." (Tony Cox comment: In multivariate regression models, measurement error can, and often does, introduce false positives and increase estimated effects sizes.)

*More generally, is it clear how criteria for individual study quality were applied to each study used in making causal determinations, and what the results were?* (See Table Annex 6-1, cf p. 6-67.) *Is it clear how the limitations of each individual study were taken into account in causally interpreting their reported associations and in making causal determinations?*

Answer: No.

*Responses from external consultants:*

- Dr. North: "**No**, I find the discussions of study quality rather superficial. It is clear that some of the recent studies have initiatives for investigating confounding factors. Much more is needed, in the context of learning how to do better estimation of the C-R relationship in the low exposure range."

*Does the ISA make clear how its causal determinations would change if evidence from associations caused by confounding, residual confounding, measurement error, or unverified modeling assumptions were excluded?*

*Responses from external consultants:*

- Dr. North: "**I did not find much discussion on confounding, measurement error, and unverified modeling assumptions.** The discussions I did find were often superficial."
- Dr. Thomas: "This seems rather speculative, absent any evidence that those studies included have failed to adequately address these possible biases." (Tony Cox comment: There is abundant evidence, including discussions by the original authors in papers such as that of Tétreault et al., that included studies did not adequately address these and other possible biases.)

*Is the biological evidence presented in the ISA to support causal determinations correctly stated, correctly interpreted, relevant for predicting effects of changes in the ozone NAAQS, and up-to-date?*

*Responses from external consultants:*

- Dr. North: "I would like to see more evidence in the ISA of the kind in the Michaudel et al. and Xu et al papers, indicating mechanisms for health damage and biomarkers or indicators for biological changes in humans at much lower exposure levels. …It will be important to get expert judgment on which cytokines, neutrophils and other indicators of inflammation are most significant for predicting irreversible damage to lung tissue such as described in the Michaudel et al. quote above. The **support in the ISA seems weak for inflammation in humans at 60-80 ppb exposure over 6.6 hours with exercise. This support is from studies done before the last review of the ozone standard**."
- Dr. Thomas: Yes.

*Does the biological evidence presented in the ISA provide well-validated scientific information suitable for predicting the effects on public health of changing NAAQS standard for ozone?*

*Answer:* No, the ISA does not address the effects on public health of changing NAAQS standard for ozone. (Tony Cox comment: Modeling the effects of changes in the PA without first providing a critical discussion of the scientific basis for such models in the ISA deprives readers of the scientific information needed to make an informed decision about whether changes in policy are needed to protect human health.)

*Responses from external consultants:*

- Dr. North: "Yes, there is some useful information, and I will consider a series of published papers reporting similar findings as well-validated evidence. But **the focus is not on assessing the effects of changing the NAAQS, as should be the case**."

- Dr. Thomas: "I believe the evidence presented in the ISA is suitable for reaching a causal interpretation of the effects of air pollution on human health. **The ISA does not address the implications of potential changes in the NAAQS**; it is my understanding that that will be addressed in the draft Policy Assessment document that I have not seen." (Tony Cox comment: My question was whether the information presented in the ISA is suitable for predicting effects on public health of potential changes in the NAAQS, not whether the ISA or the PA makes such predictions.)

*Is each of the causal determinations summarized in Table ES-1 (especially those labeled "causal relationship" or "likely to be causal relationship") the only possible causal determination conclusion that is justified by, or consistent, with current scientific evidence? Could different causal determinations be equally well justified (or better justified) by the information presented, or by the totality of current scientific evidence?*

*Responses from external consultants:*

- Dr. North: "It seems clear that at 2.5 ppm ozone exposure causes respiratory damage in mice. NIOSH scientists have declared 5 ppm "immediately dangerous to life and health" (IDLH) for short-term exposure. But for all the other categories of causality and for ambient ozone exposure levels, **I would like to see projections of what reductions in ambient levels might do to avoid adverse human health impacts**, **and similarly for welfare effects.**"
- Dr. Thomas: "**Any judgment of a "causal" or "likely to be causal" relationship is potentially subject to differences of opinion amongst experts.** It is my opinion that the various determinations summarized in Table ES-1 are well justified by the totality of the evidence, based on what I have read in sections IS.4, the supporting appendices, and my general background knowledge of the field of air pollution epidemiology. I do not claim to have read more than a portion of the ISA or to have an exhaustive knowledge of the substance matter of air pollution epidemiology, however. That said, **I would consider it highly unlikely that any other conclusions would be "equally well justified" or "better justified"** than those reached by the authors of the draft ISA. …

*Are there changes in the design, analysis, selection, or interpretation of individual studies or in the ISA's processes for interpreting and summarizing them that would improve the validity, credibility, and transparency of the ISA's scientific reasoning and conclusions?*

*Responses from external consultants:*

- Dr. North: "The overall answer is **yes on all counts**."

- Dr. Thomas: "Obviously **I would welcome wider application of the techniques of causal inference** to observational studies, along the lines of those publications I cited in my response to the draft PM2.5 PA. That said, I believe that the weight of evidence approach used by EPA to evaluate the totality of the evidence, experimental and observational, to be highly appropriate and I have no further suggestions for improvement in that process. … I do not have the comprehensive knowledge of the air pollution epidemiology literature to address whether there are key studies missing, other new evidence, or inadequately considered weaknesses."

# Dr. Mark Frampton

## Preamble

**Study Quality.** Section 4 (p. 7), regarding the assessment of study quality, does not indicate how quality assessments are used in the review. This issue was raised in the CASAC review of the PM ISA as well. The list of quality aspects that are reviewed are appropriate and complete, but nothing is provided about how these criteria are used or applied in the overall interpretation or assessment. It is not clear that the ISA consistently considers or incorporates these study quality assessments in reaching conclusions. This process should be strengthened and more fully described. The Preamble should provide details of how the study quality assessments are recorded, and of how they are considered in the development of the ISA and the PA.

Page 28, the publication referenced here, *What Constitutes an Adverse Health Effect of Air Pollution?* (ATS, 2000), should be updated: *A joint ERS/ATS policy statement: what constitutes an adverse health effect of air pollution? An analytical framework* (Eur Respir J 2017). The statement on this page that this document "…described transient decrements in lung function as adverse when accompanied by clinical symptoms", while correct, over-simplifies the issue. Transient decrements in lung function should be considered adverse in some circumstances, even in the absence of symptoms. The older ATS document provided this statement as an example of <u>one</u> of the situations where transient decrements should be considered adverse.

**Change in causality determination for short-term total mortality and cardiovascular effects.**

Section 10.3.1.4. indicates, "In instances when a "causal" or "likely to be a causal" relationship was concluded in the 2013 Ozone ISA (i.e., short-term ozone exposure and respiratory and cardiovascular effects and total mortality, and long-term ozone exposure and respiratory effects), the epidemiologic studies evaluated for those outcomes were more limited in scope and targeted towards study locations that include U.S. airsheds or airsheds that are similar to those found in the U.S., as reflected in the PECOS tool."

The rationale for limiting epi studies in these categories of causality is to emphasize the studies most relevant for policy in addressing possible changes in the NAAQS. This is reasonable for outcomes determined to be causal or likely to be causal. The problem is that, in the current ISA, for short-term total mortality and CV effects, the causality determinations were downgraded from likely to suggestive, based on the studies reviewed in the ISA, which were limited as indicated above. Part of the rationale for downgrading these causality determinations was continued limitations in the epidemiological evidence. We don't know from the ISA how many

studies were excluded from consideration based on their location, or what was the impact (if any) of those exclusions the causality determination. The question is whether that causality determination would have been downgraded had all the evidence been considered. This needs to be addressed in the ISA, with a broadening of the epi review criteria, and re-assessment of the strength of the causality relationship, for these categories of health effects.

A brief PubMed search limited to the last 5 years identified more than 40 relevant epidemiology studies examining mortality and cardiovascular disease outcomes, conducted outside of North America. The following 3 studies appeared to be of particular high quality and relevance, and were published in high quality journals:

1. Bae S, Lim YH, Kashima S, Yorifuji T, Honda Y, Kim H, Hong YC. Non-Linear Concentration-Response Relationships between Ambient Ozone and Daily Mortality. *PLoS One* 2015; 10: e0129423.

2. Bero Bedada G, Raza A, Forsberg B, Lind T, Ljungman P, Pershagen G, Bellander T. Short-term Exposure to Ozone and Mortality in Subjects With and Without Previous Cardiovascular Disease. *Epidemiology* 2016; 27: 663-669.

3. Yin P, Chen R, Wang L, Meng X, Liu C, Niu Y, Lin Z, Liu Y, Liu J, Qi J, You J, Zhou M, Kan H. Ambient Ozone Pollution and Daily Mortality: A Nationwide Study in 272 Chinese Cities. *Environ Health Perspect* 2017; 125: 117006.

**ISA, Executive Summary**

**Table ES-1**: The order of outcomes in this table should reflect the order in the document: metabolic effects are discussed after cardiovascular effects.

**Section IS.4.1** describes "Connections among health effects" in a potentially useful manner, and in a way that is not addressed in the individual appendices. However, the description on p. IS-20, of ozone effects in rats causing reductions in body temperature, BP, etc., as an example of "multisystem disruption", is somewhat confusing, because these responses do not occur in humans. The sentences following this refer to increased BP rather than decreased BP, which adds further confusion, especially since inceased BP is not an ozone response seen in the human clinical studies. This section needs to be re-thought and rewritten.

**Appendix 3**

**Figure 3-1** provides an excellent synthesis of known and suspected biological pathways mediating ozone respiratory health effects. Some suggestions for further refinement:

1. Altered heart rhythm is included here, which is obviously not strictly a respiratory response. But other nonrespiratory links are not included here, that are consequences of ANS modulation and stress responses, including systemic inflammation and metabolic processes. This seems to be an inconsistency. Would remove altered heart rhythm from this figure for consistency.

2. Impaired host defense is shown linked solely with oxidative stress, but other pathways, for which there is evidence, are likely contributing, including airway injury, morphologic airway changes, and stress responses (elevated cortisol). Consider moving this box one column to the right, ungroup from morphologic changes and allergic responses and show as one of the downstream effects.

3. The pathway indicating that adrenal effects mediate airway injury/inflammation is based on a single study in rats (Miller et al 2016b). This finding runs counter to physiologic expectations (adrenal mediated stress response would be expected to follow acute inflammation/injury, not mediate it) and there is no evidence to support that this occurs in humans. Without further confirmation in additional studies or other species, or support of this directionality in humans, suggest making this line dotted.

**P. 3-14**, last paragraph, add Frampton et al. 2015 [1] to the list of new studies of lung function effects in the range of 100-300 ppb. This study included both GSTM1 sufficient and null subjects, and showed no effects of GSTM1 gene status on lung function responses.

**P. 3-18**, Cigarette Smoking. This section summarizes the Bates et al. 2014 study as showing similar lung function responses between smokers and nonsmokers, and indicates that this finding differs from previous studies. But the smokers in the Bates study were so-called “light” smokers, on average smoking about ½ pack per day for 6 years, for a total of 3 pack-years. This likely explains the difference from prior studies, which involved subjects with greater tobacco use, and this should be noted in the summary. For example, in Frampton et al. 1997 [2], one of the studies demonstrating significantly reduced lung function effects in smokers compared with never-smokers, only smokers of at least 1 ppd for a minimum of 3 years were included. The mean pack-years of smoking was 12.8. It should also be noted that, while ozone-induced lung function decrements are attenuated in smokers, lung inflammation is not [3], and oxidative stress may actually be increased [4]. This is an example of a situation where adverse respiratory effects of ozone may be occurring in the absence of lung function changes.

**3.1.4.2.2 Animal Toxicological Studies, p. 3-23**. Symptoms by definition are self-reported, and animals are obviously unable to report symptoms. It should be more clearly pointed out here that symptoms cannot be assessed in studies of rodents. Cough, or any other change in respiratory status, when reported by an observer, is a sign or an observation, not a symptom. It is only a symptom when reported by the individual experiencing it.

**P. 3-38, line 1-2, Integrated summary**. Change "FEV1" to FVC. FEV1 is affected by changes in both volume (FVC, restrictive) and airways obstruction (FEV1/FVC).

**P. 3-46, line 8**. "These effects include sensory and pulmonary irritation…" The distinction here between "sensory" vs "pulmonary" irritation doesn't make sense. Pulmonary irritant responses have major sensory components. This phrase appears to be taken straight from the Hansen et al. 2016 abstract, but the terminology used in that abstract is not reflective of airway physiology. Sensory vagal-mediated inputs are important throughout the respiratory tract. The upper-lower airway distinction here is incorrect, and is irrelevant to the point being made in this summary. The Hansen et al. study examines pulmonary outcomes, not upper airway responses.

**Long term respiratory effects**

The first paragraph of section 3.2.1, which includes a summary of the findings from the 2013 Ozone ISA, should include the limitations and uncertainties at that time that precluded a determination of "causal" for long-term respiratory effects.

**Appendix 4.**

**Figure 4-1** provides an excellent representation of the pathways, and evidence supporting them, leading to potential cardiovascular outcomes related to ozone exposure.

**4.1.9.2, p. 4-23, 2nd bullet point**. The description of the Arjomandi 2015 study, which is a clinical study, is written as if describing an epidemiology study. This paragraph should be re-written to indicate that subjects were exposed to clean air and 2 concentrations of ozone for 4 hours, with intermittent exercise, with HRV measured before and at intervals after exposure. In such a controlled and blinded experimental exposure, the changes can reasonably be described as effects of the exposure, rather than associations.

**Table 4-4.** The study by Frampton et al. 2015 did not assess LVDP. The cardiac function outcomes were cardiac index, stroke-volume index, and left ventricular ejection time. It is perhaps worth mentioning that these measures were obtained via impedance cardiography, rather than directly or via echocardiography.

**Table 4-19.** The study of Rich et al. 2018 measured SBP as well as DBP.

**Table 4-26, Study-specific details from controlled human exposure studies of coagulation**. This table should include Frampton et al. 2017 [5], which examined a number of coagulation parameters, without significant effects.

**Table 4-29, systemic inflammation and oxidative stress markers**. Add Frampton et al. 2017 [5] here as well.

**Appendix 5, metabolic effects.**

**New determination of "likely to be causal"**. This determination is driven by the animal toxicology, which is largely limited to rodents. The animal data on glucose and insulin effects are robust. But the extrapolation of the findings to humans is in question. There appear to be no primate studies. The epidemiological evidence is sparse and inconsistent, without any evidence of adverse clinical outcomes related to metabolic effects. One human clinical study (Miller 2016a) showed no effects on insulin levels or HOMA-IR, but did find acute increases in stress hormones in response to ozone exposure. It is as yet unconfirmed. While the animal studies provide plausibility, the sparse epi and human clinical data do not justify the "likely" determination. "Suggestive" appears to be more appropriate.

One question that should be considered is whether a brief stress response, in the absence of symptoms or other consequences, constitutes an adverse health effect. This could be considered a physiological response to a variety of stimuli. For example, it can occur in response to exercise.

Nevertheless, given the potential importance of these effects for human health, and in consideration of the current epidemics of obesity and diabetes, this represents an area of urgent research need.

**5.1.4.1.** Obesity animal tox studies: Some of the studies summarized here are relevant to obesity as a risk factor, in other words, whether obesity as a subject characteristic enhances ozone responses: pulmonary, CV, or other. Descriptions of these studies should be moved to the appropriate section on risk factors. The issue being considered in this section is whether ozone alters metabolic functions including body weight, BMI, body composition, caloric intake, glucose metabolism, lipid metabolism, stress responses, etc. The mixing of these two concepts is confusing and perhaps misleading. The sentence in this paragraph starting on line 12 describes what this section should be about: "Recent toxicological studies provided some evidence that ozone may impair metabolism and affect body weight, BMI, and body composition, as well as effect [sic] caloric intake."

The new evidence for metabolic effects does support the feasibility of ozone CV effects, given the strong link between the two.

**P. 5-14**: "Acute-phase liver proteins, such as CRP, can act as sensors of liver function". This is not accurate. CRP is made in the liver, and is a marker of systemic inflammation. Its production is driven by interleukin-6, released by a variety of cells during inflammation. Although produced in the liver, it is not considered a clinically useful marker of liver function.

**5.1.5.3.4, Summary, p. 5-17.** "Elevated circulating stress hormones are consistently observed in animal models and in controlled human exposure studies after short-term ozone exposure." This should be "in a single human controlled exposure study".

The last sentence of this summary statement ("Thus, neuroendocrine stress activation is essential to the development of adverse metabolic outcomes after short-term ozone exposure.") is overly broad and not completely supported by the described (adrenalectomy) studies.

**5.1.5.4., p. 5-18. Serum lipids.** The description of the Chen et al 2016a study is unclear, and it seems incorrect. According to the Abstract, this study deals with changes in lung function and nasal inflammation among schoolchildren. Was the reference intended to be Chen 2016b?

**5.1.8., p. 5-23.** Ketone bodies as a "marker" of diabetes is not accurate. Ketone bodies are also a "marker" of starvation or consuming a ketotic (low carb) diet. It is more accurately a marker of metabolic stress or perturbation with regards to glucose utilization. It does go up with diabetic ketoacidosis and can be considered a marker of that condition, but not of diabetes in general. Transient elevation of ketone bodies does not mean a person has or will get diabetes.

**Table 5-1**. "Consistent epidemiologic evidence" is inappropriate given there is only one study supporting it.

## Appendix 6. Mortality

**Table 6-1**, under "Key Evidence". The statement is made, "Animal toxicological and controlled human exposure studies do not provide consistent evidence of potential biological pathways." Actually, the experimental animal evidence for CV effects is fairly robust and convincing. It was mostly the inconsistency in the human studies and the relative lack of CV morbidity studies that led to the change in causality determination.

**6.2.7, Summary and Causality Determination**, Long-term total mortality, p. 6-40. The following statement in this section contrasts with previous text and the overall conclusions: "There is coherence across the scientific disciplines (i.e., animal toxicology, controlled human exposure studies, and epidemiology) and biological plausibility for ozone-related cardiovascular

(Appendix 4) and respiratory (Appendix 3) endpoints, which lend some additional support to the ozone-mortality relationship." The point is made repeatedly earlier in the ISA that the clinical studies are inconsistent with regard to CV effects. This sentence needs to be reconsidered and harmonized with the rest of the document.

**Appendix 7. Other health endpoints**

**Nervous system effects**. Apparently included in this are the effects on the pulmonary irritant receptor/autonomic pathways that are well-established pulmonary effects in both animals and humans. Consideration should be given to separating this, and having this section include effects beyond the pulmonary irritant response loop, perhaps limiting it to brain, cognitive, and behavioral effects. Otherwise this category is causal based on the known local pulmonary neurological effects.

**P. 7-42, line 21**, "reproductive" effects should presumably be "nervous system" effects here.

**References**

1. Frampton M, Pietropaoli A, Dentler M, Chalupa D, Little E, Stewart J, et al. Cardiovascular effects of ozone in healthy subjects with and without deletion of glutathione-S-transferase M1. Inhal Toxicol. 2015;27:113-9.

2. Frampton MW, Morrow PE, Torres A, Cox C, Voter KZ, Utell MJ. Ozone responsiveness in smokers and nonsmokers. Am J Respir Crit Care Med. 1997;155:116-21.

3. Torres A, Utell MJ, Morrow PE, Voter KZ, Whitin JC, Cox C, et al. Airway inflammation in smokers and nonsmokers with varying responsiveness to ozone. Am J Respir Crit Care Med. 1997;156:728-36.

4. Voter KZ, Whitin JC, Torres A, Morrow PE, Cox C, Tsai Y, et al. Ozone exposure and the production of reactive oxygen species by bronchoalveolar cells in humans. Inhal Toxicol. 2001;13:465-83.

5. Frampton M, Balmes J, Bromberg P, Stark P, Arjomandi M, Hazucha M, et al. Multicenter Ozone Study in oldEr Subjects (MOSES). Part 1. Health Effects Institute Research Report 2017;No. 192.

# Dr. Sabine Lange

My main comment topics are discussed at the beginning of this document, with the details for each comment described after these summaries, followed by the expert consultant responses to my questions.

A reference list can be found at the bottom of this document for those studies that are not referenced in the ozone ISA.

***Charge Questions:*** *Please comment on the identification, evaluation and characterization of the available scientific evidence from epidemiologic, controlled human exposure, toxicological and associated human exposure and atmospheric sciences studies and the application of information from these studies to inform causality determinations for human health outcomes.*

- *Appendices 3 - 7 present assessments of the health effects associated with short-term and long-term exposure to ozone. The discussion is organized by exposure duration, broad health effects (e.g., asthma, ischemic heart disease, etc.), and scientific discipline. Please comment on the characterization of the evidence within these chapters.*
- *Please comment on the portrayal and discussion of the biological plausibility evidence presented in Appendices 3-7 and the extent to which: (1) the organization adequately captures the current state of the science with respect to potential pathways by which ozone could impart health effects, and (2) as currently constructed, inform causality determinations.*

<u>Study Quality</u>

The EPA has improved their systematic review and study quality assessment. However, the study quality review needs further development.

- Only certain studies were included in the quality analysis – while the EPA notes that those from causal and likely causal designations were included, not all the studies from those appendices had quality evaluations in HAWC (for example, for the long-term metabolic epidemiology studies – only one of the 6 studies cited in causal designation Table 5-4 had a study quality evaluation listed in HAWC).
- It is not clear how the study quality assessment is integrated into the main text and how it informs the conclusions based on the evidence.
- Given that the causality determinations explicitly state that there is evidence from a "high quality study" it is still not clear which studies the EPA considers to be "high quality".

- Chance, bias, and confounding are all potential reasons for a study to observe an association between two variables (Zaccai, 2004) and therefore should be more explicitly considered when presenting and discussing study results.
    - More factors than just copollutants should be considered as important confounders in the referenced epidemiology studies.
    - Results that are not statistically significant should be indicated as such in the ISA discussion. If there is a reason why statistical significance may not have been achieved (e.g., low sample size), this should be included in the discussion of the study results. The general conclusion from the expert consultants was that statistical significance does need to be given some consideration, but there are also other factors such as patterns in the epidemiology data that also should be factored into the conclusions that are drawn.

<u>Accuracy of Presentation</u>

The EPA should provide a balanced summary of the study results for each health endpoint. Adequately communicating available positive, negative, and null results provides useful information for further documents in the ozone NAAQS review.

- In section summaries, divergent results should not be ignored, but rather should be included in a more nuanced summary of results. For example, the Arjomandi et al. 2018 study did not find an association between GSTM1 phenotype and inflammation. However, in the summary section for respiratory effects in healthy populations this divergent finding was not included or even intimated: the EPA stated that "Recent studies are consistent with previous findings and expand on observed interindividual variability in inflammatory responses, providing additional evidence that GSTM1-null individuals are more susceptible to ozone-related inflammatory responses."
- Further, information summarized from one section to another should maintain the accuracy and nuance of the underlying data. For example, in sections 4.1.16 and 6.2.4.1, the EPA states "Specifically, the evidence from controlled human exposure studies provided support for increased decrements in FEV1 and greater inflammatory responses to ozone in individuals with asthma than in healthy individuals without a history of asthma." This is not the case, the respiratory chapter addresses this point at length, and states that people with asthma are not less sensitive than people without asthma for lung function effects.
- The EPA should provide accurate study information as well as study conclusions. The study results presented in the metabolic chapter are particularly error-prone and need to be reviewed carefully. For example,
    - In section 5.1.4 (overweight and obesity) the EPA states that "Ozone exposure caused males on the control and high-fat diets to eat statistically significantly more food and trended toward statistically significant increases on high fructose

diet (Gordon et al., 2016)." The exposure information is incorrect in the summary of this study (animals exposed one time per week, not 4 times per week). Also, the animals eat more, but they don't gain more weight - is ozone increasing their metabolism, such that they don't gain weight with increased caloric intake?

- In section 5.1.5.1 (Other effects, inflammation), the EPA states that "Obesity-prone mice (adult male KK mice) were exposed to ozone for 13 consecutive weekdays [4 hours/day; Zhong et al. (2016)]." However, Table 5-10 says that the exposure was 3 consecutive days.
- In section 5.1.5.4 (Other effects, serum lipids), the data presented for the Gordon et al. 2016 study (should be 2016b) is inaccurate. The EPA states that "The effect of high-fat and high-fructose diets was tested in male brown Norway rats" – the study was done in male and female rats; "With ozone exposure (0.8 ppm ozone, 4 days/week for 3 weeks)," – exposure was 0.8 ppm ozone, 5 hrs/day, 1 day/week for 4 weeks (subacute exposure), or a single 0.8 ppm exposure for 5 hrs (acute); "Females were refractory to change." – the abstract of the paper says "Female rats appeared to be more affected than males to O-3 regardless of diet."
- In section 5.2.5 (metabolic syndrome and type II diabetes), the Jerrett et al. 2017 effect estimates are incorrect (presented are 1.28; 95% CI: 1.06, 1.55 and 1.20; 95% CI: 0.96, 1.50 with NO2 adjustment – should be 1.18 (1.04, 1.34) and 1.13 (0.97, 1.31) with NO2 adjustment.

<u>Dose Assessment and Concordance</u>

The EPA should appropriately compare animal to human ozone doses when extrapolating animal exposures to potential human risks.

- The PECOS statement for experimental studies in Appendix 3 (and on pp 3-19 and 3-26) notes that resting rats exposed to 2 ppm have an equivalent ozone deposition as exercising humans, citing Hatch et al. 1994. The EPA should further discuss that there is a direct relationship between resting human and resting rat ozone doses, and that a human with a ventilation rate that is five times higher than resting will have a 5 times higher dose. This should be correctly noted and Hatch et al. 2013 and McCant et al. 2018 (describes this misconception) should be cited.
- The EPA should also consider dose in their biological plausibility discussion, in particular the contrast between known personal exposure doses (which are typically very low because people spend most of their time indoors and ozone is largely an outdoor pollutant) and the concentrations at which the observed effects occurred.
- My summary from the opinions provided by the expert consultants on the questions of animal dosimetry is: Given that the causality determination for metabolic effects of ozone exposure is mostly derived from animal toxicological studies, it is appropriate for the

EPA to more thoroughly discuss the dosimetric similarities and differences between animals and humans, beyond the simplistic reference to Hatch et al. 1994.

Clarity of Presentation

The EPA should clearly present the findings in each of the ISA sections, and should provide an accurate and balanced summary of results.

- When discussing the results from all studies, and particularly CHE studies, it is important to include the exposure duration (e.g. on pg 3-26 when discussing concentrations at which airway hyper-responsiveness has been observed) and the exercise level of the participants (e.g. in the integrated synthesis when discussing concentrations that could generate adverse effects in healthy adults).
- For the discussions addressing pre-existing conditions, the EPA should specifically note data that provides information on responsiveness of people with the condition to people without (because this goes directly to potentially sensitive subpopulations). For example,
    - In section 3.1.5.3 when the EPA discusses CHE studies, the EPA notes that in Horstman et al. 1995 they show that there are more wheeze symptoms in people with asthma exposed to ozone - is this compared to people without asthma, or ozone just increases wheeze in general?
    - In section 3.1.6.2 the EPA describes the data showing whether there may be more sensitivity to respiratory effects of ozone of people who are obese or who have metabolic syndrome. However, in describing the study results, particularly of Ying 2016, Zhong, 2016, or Gordon 2016b, the EPA does not note whether there was a greater (or different) inflammatory response to ozone in the obese/metabolic syndrome animals versus lean/normal animals. Because this is the purpose of this section, these pieces of information should be included.
    - The EPA includes sections about respiratory effects in pregnancy (3.2.4.7) and in populations with metabolic syndrome (3.2.4.8). Is the purpose of these sections to show that there is an increased response to ozone in these populations? If so, then the EPA should specifically provide information and discuss whether the data shows that these groups are more sensitive. As it stands, this conclusion is not clear.
    - Is section 5.1.4 (overweight and obesity) intended to discuss the impact of ozone on obesity, or the effects of ozone on overweight/obese individuals? If it is the latter this should be part of the other chapters in the sections discussing sensitive subpopulations.
- If possible, the EPA should avoid making statements that address an unlikely conclusion, but that avoid addressing the conclusion of interest. For example:

- In section 3.1.5.4 (lung function): "it was concluded that individuals with asthma were at least as sensitive to acute effects of ozone as healthy individuals." The conclusion of interest is whether people with asthma are *more* sensitive or not.
- "However, despite limited evidence demonstrating increased sensitivity to ozone in individuals with asthma compared to those without asthma, there is consistent evidence that asthmatic individuals experience lung function decrements in response to acute ozone exposures." I don't think that there is a reasonable hypothesis that people with asthma would not experience ozone-induced lung function decrements.

- The EPA's underlying concern about people with asthma is perhaps not that they will have an increased innate response to ozone exposure (they do not seem to have greater lung function decrements, inflammation or airway hyperresponsiveness), but rather that they likely have less of a buffer against adverse effects. This is an important argument that EPA should emphasize when discussing the respiratory effects of ozone exposure on people with asthma.
- The EPA has described the exercise level in key CHE studies such as Schelegle et al. 2009 as a slow walking pace, but the authors of that study note that "This protocol contains six 50-minute exercise periods with minute ventilation maintained at 8 L/min/L of FVC (VE of approximately 40 L/min). As noted by Folinsbee and colleagues [Folinsbee et al. 1988] and McDonnell and colleagues [McDonnell et al. 1991], this level of exertion was ''intended to simulate work performed during a day of heavy to severe manual labor in outdoor laborers.'" This discrepancy in description should be clarified
- In section 4.1.8 (blood pressure changes and hypertension) ED visits and HAs, the EPA puts the findings into the context of the mean ozone concentrations measured in those areas. This is an attempt, I think, to understand the results in the context of a dose-response. This type of discussion is very helpful and should be included elsewhere.
- When the EPA states that there is little evidence for ozone impacting a particular endpoint, they should clarify whether there is little evidence because the studies haven't been done, or that the available studies do not show an association. For example, in section 4.2.2 (biological plausibility) the EPA notes that "However, considerable uncertainty remains in how long-term ozone exposure may lead to mortality given that there is little epidemiologic evidence of an association between long-term exposure to ozone and other cardiovascular endpoints such as IHD, stroke, or thromboembolic disease." Does this mean that studies have not been done, or that studies have been done that have not shown associations?
- In section 5.1.5.4 (Other effects, serum lipids), the EPA states in multiple locations that in an animal study, certain groups were "refractory to change". For example, "all other endpoints (HDL, LDL, and total cholesterol), ages, and doses (0.25 ppm ozone) were refractory to change (Bass et al., 2013)." Refractory means resistant and suggests that the endpoint would have changed but there was some active mechanism that kept them from changing. These endpoints weren't refractory, they just didn't change.

Consistency of Results & Reporting

The ISA would be strengthened by more justification of decisions in the face of conflicting evidence. An example of inconsistent (or seemingly inconsistent) results comes from section 5.2.3 (glucose and insulin homeostasis) where evidence is presented from three studies (Miller et al. 2016b, Gordon et al. 2013, Bass et al. 2013). Miller, Gordon, and Bass all came from essentially the same set of authors and tested effects of long-term ozone exposure in male rats. But they show different effects: Bass showed no change in fasting glucose with subchronic exposure, but Miller did; Miller showed decreased baseline insulin in subchronic exposed adult animals, but Gordon showed no change in adult exposed animals, and increases in insulin in senescent exposed animals. The EPA should speak to whether there are patterns in these results, or if the differences are spurious or show strain differences.

Consistency in EPA's reporting and interpretation of results is also important. For example, in section 4.1.17 (causality determination) the EPA states that "Studies from Europe, Canada and the U.S., several of which analyzed a large number of events per day in multiple cities, consistently reported null or only small positive effect estimates (i.e., OR ≤ 1.02) in analyses of MI, including for STEMI and NSTEMI (Section 4.1.5.1)." This is the only section where I have seen the magnitude of the association considered by EPA. Is there a reason why a small magnitude effect for this endpoint would be more important that a small magnitude effect for other endpoints? If the EPA is going to consider magnitude of effect for these studies, they should be clear as to why, and whether this is also a relevant consideration for other endpoints.

Applicability of Results from Animal Studies

Dose-responsiveness of effects of ozone exposure in experimental studies can be used to identify relevant biological plausibility pathways and exposure-specific responses, and so should be further discussed in those sections. In particular establishing no-effect and low-effect concentrations for endpoints such as long-term ozone exposure and lung function development would ease the extrapolation to effects in humans at ambient concentrations.

In addition, information about the comparability of animal models to human disease are useful in extrapolating results from animal studies – such as information about how good of a model allergic airway disease in mice is compared to humans. Even more important is information allowing the interpretation of *ex vivo* studies, such as experiments in isolated, perfused hearts (section 4.1.4).

Shape of the C-R Function

As was discussed in the CASAC's responses to the PM ISA and PA, errors and heterogeneity in epidemiology study variables can affect the apparent shape of the concentration-response (C-R)

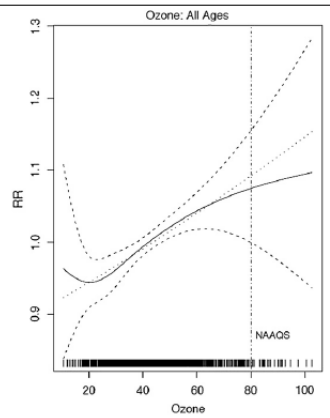
relationship and can obscure thresholds. Evidence for this has been provided by many peer-reviewed publications (Brauer et al., 2002; Cox, 2018; Lipfert and Wyzga, 1996; Rhomberg et al., 2011; Watt et al., 1995; Yoshimura, 1990) and notably by the EPA in the ISA preamble (US EPA 2015, Section 6c, pg. 29):

> "Various sources of variability and uncertainty, such as low data density in the lower concentration range, possible influence of exposure measurement error, and variability among individuals with respect to air pollution health effects, tend to smooth and "linearize" the concentration-response function and thus can obscure the existence of a threshold or nonlinear relationship. Because individual thresholds vary from person-to-person due to individual differences such as genetic differences or pre-existing disease conditions (and even can vary from one time to another for a given person), it can be difficult to demonstrate that a threshold exists in a population study. These sources of variability and uncertainty may explain why the available human data at ambient concentrations for some environmental pollutants (e.g., PM, O3, Pb, environmental tobacco smoke, radiation) do not exhibit population-level thresholds for cancer or noncancer health effects, even though likely mechanisms include nonlinear processes for some key events."

The problem described here is not whether a threshold in the data may exist, but rather that even if it does exist, epidemiology studies may not be capable of definitively identifying the threshold. To address this concern the EPA should explicitly acknowledge in the ozone ISA that variability and error in the variables can linearize C-R functions and obscure thresholds, and this acknowledgement should be included in those places where the EPA concludes that the relationship between ozone and a health effect is linear and has no threshold. I also recommend that the EPA begin to apply methods (and encourage the epidemiological community to apply methods) to address this particular concern, including errors-in-variables methods. If possible, the EPA should include these types of adjustments when applying the epidemiology C-R functions to their risk assessments.

In section 6.1.7 (shape of the C-R function), the EPA states that in the previous ISA the available studies showed no evidence of a deviation from linearity or the presence of a threshold for short-term ozone-mortality relationships. "However, it is important to note that the examination of the ozone-mortality C-R relationship is complicated by previously identified city-to-city and regional heterogeneity in ozone-mortality risk estimates (U.S. EPA, 2013a). Recent studies continue to provide evidence of a linear C-R relationship with no evidence of a threshold below which mortality effects do not occur along the distribution of ozone concentrations observed within the U.S." The EPA should provide information here noting whether the new studies address the consideration of city-to-city or regional heterogeneity that were concerns before, or if this is still an issue. If it is, the EPA should state it as such.

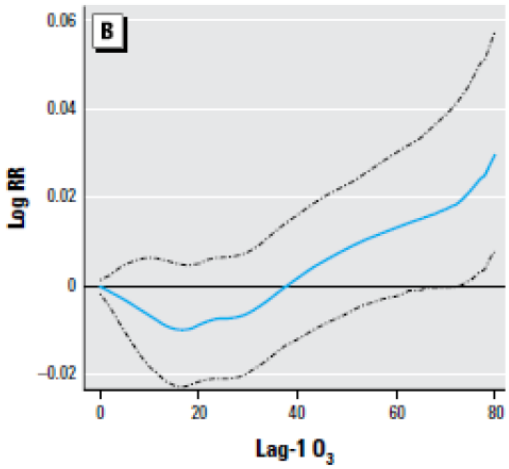
In addition, some of the plots that are presented by the EPA do not look linear and do appear to have a threshold, such as the Silverman plot (Figure 3-9), the Moolgavkar plot (Figure 6-6) and the Di plot (Figure 6-7). If the EPA thinks that there is so much uncertainty at the lower ends of these curves that we cannot trust the apparent U shape, then we also cannot trust that the shape is linear, and no conclusions should be drawn. For Silverman (Figure 3-9) there is still a lot of data at the point where the curve bottoms out (about 20 ppb), so the uncertainty cannot be all about data availability.




Note: The average of 0 day and 1 day lagged 8-hour daily max ozone was used in a two-pollutant model with $PM_{2.5}$ lag 0−1, adjusting for temporal trends, day of the week, and immediate and delayed weather effects. The solid lines are smoothed-fit data, with long broken lines indicating 95% confidence bands. The density of lines at the bottom of the figure indicates sample size. The NAAQS line indicated in the figure is reflective of a previous standard set in 1997. The form of this NAAQS was the 3-year avg of annual 4th highest daily max 8-hour concentrations.

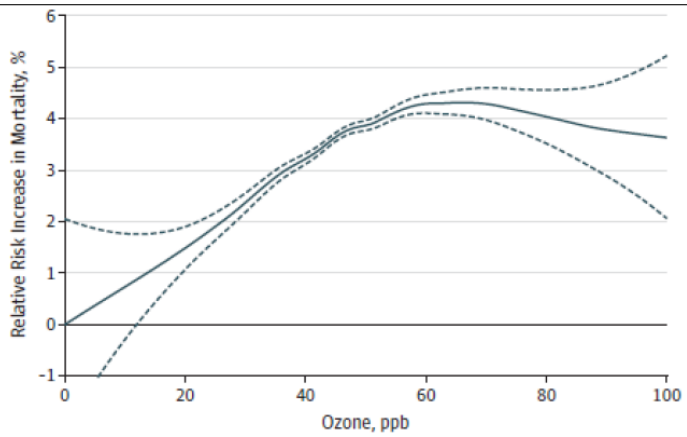
Source: Permission pending, Silverman and Ito (2010).

**Figure 3-9 Estimated relative risks (RRs) of asthma hospital admissions for 8-hour daily max ozone concentrations at lag 0-1 allowing for possible nonlinear relationships using natural splines.**




Source: Permission pending, Moolgavkar et al. (2013).

**Figure 6-6 Flexible concentration-response relationship for short-term ozone exposure and mortality at lag 1 for 24-hour avg ozone concentrations adjusted by size of the bootstrap sample (size of the bootstrap (*d*) = 4).**




Source: Permission pending, Di et al. (2017a).

**Figure 6-7 Percent increase in mortality for ozone in a two-pollutant model with $PM_{2.5}$ using penalized splines for both pollutants at lag 0−1 days in the warm season (April−September).**

Interpretation of Study Results

The ISA should address the adversity and clinical significance of the presented health effects. For example:

- In section 5.1.3 (glucose and insulin homeostasis) the EPA presents epidemiology results that demonstrate 0.5-3% chances in fasting glucose concentration. It seems that the normal range is ~70-100 mg/dL, which is an up to ~40% variability, so what is the clinical significance of these types of findings?
- In section 5.1.4 (overweight and obesity) the EPA states that "Reduction in body weight gain has also been noted in healthy nondiabetic rats exposed to ozone". Is a reduction in body weight gain in healthy and obese animals an adverse effect?

In addition, inclusion of the significance of relationships between different factors identified in epidemiology studies would help clarify the conclusions that can be drawn. For example:

- In section 3.2.4.1 the EPA discusses findings from the Children's Health Study noting that there is a relationship between ozone exposure, new onset asthma, and particular genetic variants. What is the significance of the relationship between genetic variants and new-onset asthma and ozone risk in the CHS study? What do those genes do? It is difficult to interpret genetic variant information in the absence of contextual information when determining the risk of new onset asthma with ozone exposure.
- In section 4.1.16 (effect modification) for pre-existing disease the EPA presents information about ozone-associated changes in HR or BP in people with mood disorders. What is the significance of mood disorders for ozone-associated changes in HR or BP?
- In section 5.1.5.1 (Other effects, inflammation), the EPA states that the Zhong et al. 2016 study showed an increase in inflammatory mediators in epididymal adipose. The EPA should provide information about the significance of increase inflammation in epididymal adipose, as well as whether inflammation was seen in other visceral adipose tissues.

It would be helpful if the EPA provided information about how short-term exposures impact long-term effects, how long-term exposures impact short term effects, and how effects of short- and long-term exposure are separated. For example:

- In section 4.1.8 (blood pressure changes and hypertension) the EPA notes that hypertension is a chronic condition that develops over a period of years. Therefore, how is an ozone exposure of several hours contributing to hypertension emergency department (ED) visits or hospital admissions (HAs)? Are the study authors (and the EPA) postulating that ozone concentrations trigger an immediate change in BP that sends the person with hypertension to the hospital? Sort of like an asthma attack? The EPA should clarify how ozone is expected to contribute to this endpoint.

Completeness of Study Information

The EPA should ensure that all relevant information is included in the study figures or tables. For example:

- In section 3.1.10.1 (copollutant confounding), the EPA notes that they provide study-specific details in the tables in Section 3.3. However, the information in those tables do not include the effect estimates for the copollutant models, only for the single pollutant models. The EPA should include the copolluant effect estimates in these tables, or in the text or figures of this section. The latter would be preferable, because of the importance of considering copollutant confounding. Similarly for the results that consider confounding by aeroallergens.
- Table 5-7 does not include all of the information about the Ramot et al 2015 study – only one rat strain is included and not the 8 that were tested, and only one of the 3 ozone doses is included.

Causality Determinations

For the short-term ozone effects on metabolic endpoints, it is difficult to tell if the causality determination is warranted, because the presented data is not always accurate (see "accuracy in presentation" section of these comments). In addition, there is no consistent direction of effect presented – if biomarkers change in different directions in different experiments, does that matter for the EPA's causality determination? In addition, there is one CHE study presented that corroborates the animal studies, and one epidemiology study. However, multiple epidemiology studies are presented that have null results.

For the long-term ozone effects on metabolic endpoints there is again a problem of accuracy of reported results. Overall for this causality designation, there is limited epidemiology evidence (and that has issues, or associations are lost with copollutants, or copollutants aren't assessed, study quality was only assessed in one of the six epidemiology studies cited in Table 5-4, and there at least one of the study quality details was wrong). The animal evidence is not always summarized correctly and shows somewhat inconsistent results. It does consistently show no effects at lower ozone concentrations (0.25 ppm), and all 3 of the cited animal studies were conducted in whole or in part by the same group of authors. Further analyses are provided in the detailed section of these comments.

For ozone and reproductive effects, the effects of ozone on male reproduction are based on little data (inconsistent epidemiology studies, one animal study), and additionally the EPA states for female reproduction that "In conclusion, results from epidemiologic studies are mixed, with benefits and detriments to female reproductive function with ozone exposures, while toxicological studies show limited evidence of effects on successful completion of pregnancy."

Therefore, it is not clear why the EPA has designated male and female reproduction as "suggestive of causality".

Biological Plausibility

I encourage the EPA to present both positive and negative studies when presenting biological plausibility pathways, as well as information about exposure concentrations.

In section 5.1.2 (biological plausibility) the EPA needs to distinguish better between short-term and long-term effects of ozone on metabolism. It seems that they are ascribing short-term ozone exposures to diabetes development? In a comparable situation with asthma, the EPA does not attribute short-term ozone exposure to asthma development, but rather to exacerbation. The EPA states that "All of these upstream factors of autonomic activation and homeostatic imbalance can contribute to an animal model or humans being at a greater risk for developing metabolic syndrome or diabetes with ozone exposure." It is this last piece that is tricky to separate from chronic exposure effects. These axes are self-regulating, and although a single acute exposure may unbalance them, there is no evidence presented that this is irreversible. It seems like only unbalancing the system over and over again (chronic exposure) would predispose to metabolic disease.

**Further Detailed Comments**

Study Quality

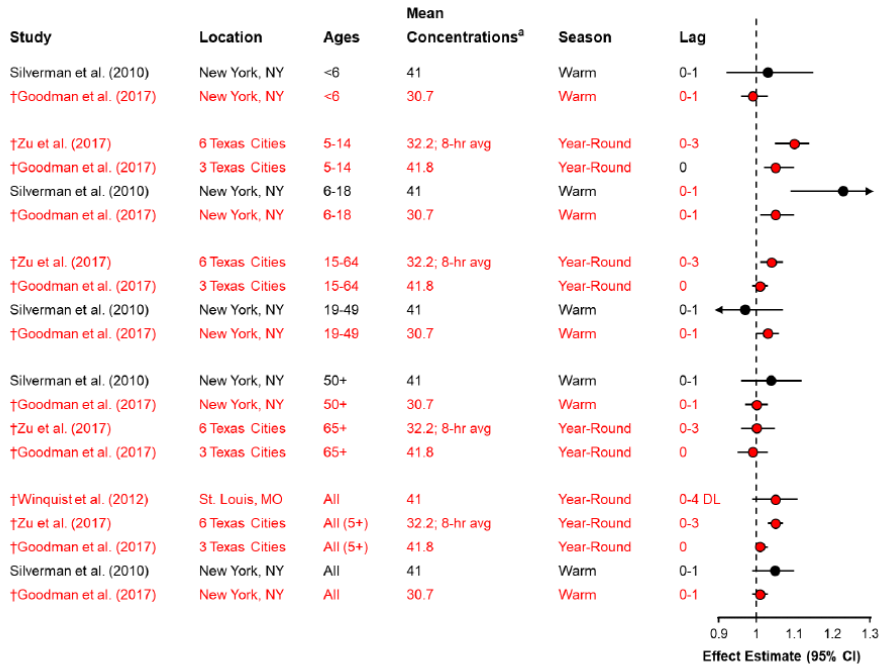
- Based on the information in Appendix 10, the EPA has done a better job (compared to the PM ISA) describing the methods used for finding studies, screening them, and including or excluding them. I still have concerns about the study quality review, however. The EPA does note that there is a study quality review and they share the review criteria, and reference HAWC where the study quality assessments can be found. However, it is not clear how that information is integrated into the ISA, and I can't find the guidance text and prompting questions that EPA refers to on pg 10-21. In addition, the EPA states that they do not use individual study quality to assess results, but rather considers the quality of the literature as a whole. There are a few potential problems with this: 1) The causal determinations state that there is evidence from a "high-quality study" - without identifying what this study is, how can a reader tell if the causality determination is actually based on an objectively-determinable high quality study, or if in fact the literature as a whole has all of the pieces that could make a high quality study, even if any one study does not have all of those pieces? 2) The study quality does matter when drawing conclusions - there are certain kinds of conclusions that can only be drawn based on certain study types and quality, and if you are weighing contradicting evidence,

then the weight should be placed more heavily on the higher quality study. If this is not made clear, then how does the reader know how to weight the studies (or how the studies were weighted by EPA?).

- It is not clear how the EPA chose to include studies in their HAWC study quality evaluation. For example, for section 5.2.5 (long-term ozone exposure and metabolic effects; metabolic syndrome and type II diabetes), only one of the three cited epidemiology studies had study quality evaluations in HAWC (Jerret et al. 2017, not Renzi et al. 2017, or Yang et al. 2018). For section 5.2.6 (metabolic syndrome and mortality) neither of the two cited epidemiology studies had study quality evaluations (Turner et al. 2016, Cruise et al. 2015).
- In the PECOS statement for experimental studies, the comparison group is stated to be the person themselves or an appropriate comparison group for CHE studies, or in animal tox studies to be a group exposed to a negative control. Negative controls are crucially important in these types of studies and should be required for both animal and human studies. In the inflammation section (3.1.4.3) the EPA notes the importance of the filtered air control by discussing exercise-induced effects, but then discusses results from studies without this control as if it is of the same value as other studies with the appropriate controls.
- In the integrated summary sections for respiratory effects, there is no distinction between those sections with lots of data (such as lung function), rather than those with few (such as symptoms in people with asthma - only one CHE cited, with only single-city epi studies and one panel study). Both are presented as having consistent evidence with no note about the strength of the overall database.
- In section 3.1.8 which discusses respiratory ED visits and HAs, study quality considerations are only discussed for one study, Winquist 2012, which did not show an association between ozone and all ages respiratory HAs (the EPA noted that only a single monitor was used for the exposure assessment). However, the same study was used without caveat as evidence for ED visits. The EPA needs to consider study quality for all studies, not just those that provide disparate information from their hypothesized effect.
- In section 6.1.6.1 (copollutant confounding), the EPA states that "The increase in the widths of the confidence intervals observed in these studies is consistent with a decrease in precision due to the limited data available to conduct copollutant analyses due to the PM sampling schedule." The EPA should provide information about how limited the data has really become. Often these NMMAPs studies have millions of data points, so dividing by 6 (for a ⅙ PM sampling schedule) still provides hundreds of thousands of data points for the analysis. This would seem to provide adequate power for the analysis.
- In section 7.1 (reproductive and developmental effects), the EPA states that "Well-designed studies that consider sources of bias, including potential confounding by copollutant exposures, are emphasized." I do not see this statement made in the other chapters. Are the well-designed studies emphasized in other chapters as well?

Accuracy of Presentation

- In this document the EPA presents epidemiology study results with different averaging times on a scale, to allow direct comparison of results. However, using a simple conversion does not capture the uncertainty that results when "converting" one averaging time to another, and may bias the resulting concentration estimates (Lange et al. 2018). This has also been demonstrated by Anderson and Bell (2010), who found that interconversion amongst different averaging time metrics of ozone introduces uncertainty, and the ratios between the averaging times could differ across communities, as well as within communities by temperature, season, and long-term ozone concentration. In addition, the Lange 2018 study shows that the 8-hr maximum and 24-hr average are not substantively correlated, so there should not be interconversion between the two.
- In section 3.1.4.4 the EPA notes that in the large and well-conducted MOSES study (Arjomandi et al. 2018) there is no association between GSTM1 phenotype and inflammation. However, this finding is diminished by the statement at the end of the section "Recent studies provide some further evidence that GSTM1-null individuals are more susceptible to ozone-related inflammatory responses, although the evidence is not entirely consistent." Then in the summary section for respiratory effects in healthy populations, this finding is reduced to "Recent studies are consistent with previous findings and expand on observed interindividual variability in inflammatory responses, providing additional evidence that GSTM1-null individuals are more susceptible to ozone-related inflammatory responses.", where the discrepancy is not mentioned at all. A similar statement is made in section 3.1.11 describing the short-term ozone exposure respiratory effects causality conclusion. This sequence demonstrates my concern with how EPA summarizes evidence, by just dropping divergent results, and even in the face of conflicting evidence, summarizing results as "consistent".
- In section 3.1.5.1 the EPA states that "Recent studies expand the existing evidence base and provide consistent evidence of an association between ozone and hospital admissions for asthma (Figure 3-4)." However, that figure shows mostly just a few positive associations for children 5-18, and a lot of null associations. This is not consistent evidence.
- In section 3.1.5.4 (lung function) the EPA presents data from Horstman et al. 1995 that summarizes the findings that FEV1 increases with individuals who use bronchodilators. Later in this section the EPA notes that Bertoli et al. 2013 had similar findings to Horstman et al. 1995, including that FEV1 decrements increased with a lack of inhaled corticosteroid treatment. This is the opposite result from Horstman 1995, not a consistent result.




DL = distributed lag.

Note: †Studies published since the 2013 Ozone ISA. Black text = studies included in the 2013 Ozone ISA.

[a]Mean concentrations reported in ppb and are for 8-hour daily max averaging times unless otherwise noted.

Results standardized to a 15-ppb increase in 24-hour avg, 20-ppb increase in 8-hour daily max, or 25-ppb increase in 1-hour daily max ozone concentrations. Corresponding quantitative results are reported in Table 3-5.

**Figure 3-4 Summary of associations from studies of short-term ozone exposures and hospital admissions for asthma for a standardized increase in ozone concentrations.**

- In section 3.1.6.3 describing respiratory responses in those with pre-existing cardiovascular disease, the EPA makes a fairly blanket statement that ozone enhances respiratory inflammation, pulmonary damage, etc, more in animals with hypertension/CVD disease compared to normal animals. But the summary of the studies shows that there is conflicting evidence of this - one whole set of studies shows lesser effects in the sensitive animals at lower concentrations of ozone, and another set shows that the age of the animal mediates the responses. More nuance needs to be applied to the summary of this information.
- In section 3.1.9 for respiratory mortality, the EPA concludes that there is a consistent positive association between ozone and respiratory mortality. Most of the evidence is from the 2013 ISA, with only a few studies adding to it - a multi city study showing an association, and a single-city study not showing an association. The evidence in the 2013 document is underwhelming - only about 1 page of discussion about 5 studies. Based on information from Table 6-5 of the 2013 ISA, of the presented studies: Bell 2005 was not statistically significant; the effect estimates in Katsouyanni 2009 were only statistically

significant in Canada in the summer (not the US or Europe in all-year or summer; or Canada in all-year); Klemm as noted in the ISA is negative, and Vanos 2014 is statistically significant positive (with much higher effect estimates than any other study). The Bell 2005 study was a meta-analysis and found non-statistically significant effects. This evidence is not consistent with the EPA's conclusions about this endpoint.

- In section 3.1.11 describing the causality conclusion for short-term ozone exposure on respiratory effects, the EPA notes that "The strongest single-day associations were generally observed with ozone concentrations on the same day as the outcome,". However, according to the EPAs Table 3-1 this is not true, the strongest lags were distributed across the four days prior to the event, depending on the study.
- In section 4.1.11 (systemic inflammation and oxidative stress) the EPA concludes for panel studies that "Altogether, these epidemiologic panel studies provide evidence that short-term ozone exposure is associated with increased inflammatory responses." However, there is about an equal number of studies showing effects as no effects, which does not provide strong evidence for this conclusion.
- In section 4.1.16 (effect modification) for pre-existing disease, the EPA summarizes the respiratory effects evidence. It is not clear why this summary is included in the CVD section, it belongs in the respiratory section. In addition, the EPA states that "Specifically, the evidence from controlled human exposure studies provided support for increased decrements in FEV1 and greater inflammatory responses to ozone in individuals with asthma than in healthy individuals without a history of asthma." This is not the case, the document states that people with asthma are not less sensitive than people without asthma for lung function effects. This same mis-statement is also made in section 6.2.4.1 (long-term mortality, pre-existing disease).
- In section 4.1.17 (causality determination) the EPA states that "Animal toxicological studies published since the 2013 Ozone ISA provide generally consistent evidence for impaired cardiac function and endothelial dysfunction" However, the section describing animal toxicological studies of cardiac function (4.1.4.3) present as many studies/results showing no effect, as show an effect. This is not consistent evidence of impaired cardiac function.
- For appendix 5 (metabolic effects), because this is a new endpoint, it is particularly important for the EPA to provide balanced information about negative, null, and positive evidence.
- In section 5.1.3 (glucose and insulin homeostasis) The EPA states that "Recent epidemiologic studies provide some evidence of associations between short-term ozone exposure and these endpoints (Table 5-5)." However, the evidence presented is mostly null, or in opposite directions (one study showing increasing fasting glucose, another showing decreasing fasting glucose). Also, what is the significance of the presented changes in fasting glucose (e.g. 0.5-3% changes)? It seems that the normal range is something like 65-90, which is an up to 75% variability. It also seems that these factors would be highly affected by diet. Is diet controlled for in any way in these epi studies?

- In section 5.1.5.1 (Other effects, inflammation), the EPA states that "Obesity-prone mice (adult male KK mice) were exposed to ozone for 13 consecutive weekdays [4 hours/day; Zhong et al. (2016)]." However, Table 5-10 says that the exposure was 3 consecutive days. The EPA needs to carefully check all the study details in this appendix to ensure accurate reporting of results, as well as making sure that all of the details are complete (e.g. some study results present the number of hours per day for ozone exposure, and others do not).
- At the end of section 5.1.5.3 (Other effects, endocrine hormones), the EPA summarizes "Removal of the neuroendocrine input by surgically removing the adrenal glands removes the neuroendocrine stress activation, ameliorates the stress hormone response and attenuates glucose intolerance and other factors that contribute to metabolic syndrome in rodents exposed to ozone." As discussed by EPA in their study summary, adrenectomy only ameliorates some of the effects, so this is an overly general summary statement that does not correctly capture the data. It also does not capture the opposite effects of ozone on leptin in healthy versus diseased animals, which should be addressed.
- In section 5.1.5.4 (Other effects, serum lipids), the data presented for the Gordon et al. 2016 study (should be 2016b) is inaccurate. Here is the study summary that is presented on pg 5-20: "The effect of high-fat and high-fructose diets was tested in male brown Norway rats. With ozone exposure (0.8 ppm ozone, 4 days/week for 3 weeks), there was significantly decreased serum cholesterol in animals on control diet, an effect which was ameliorated with high-fat or high-fructose diets (Gordon et al., 2016). In fact, ozone induced statistically significantly increased cholesterol in male animals on the high-fat diet versus high fat filter air controls. Serum triglycerides were significantly increased in ozone-exposed male rodents on the control or high-fat diets versus filter air controls. Females were refractory to change."
    - "The effect of high-fat and high-fructose diets was tested in male brown Norway rats" – the study was done in male and female rats
    - "With ozone exposure (0.8 ppm ozone, 4 days/week for 3 weeks)," – exposure was 0.8 ppm ozone, 5 hrs/day, 1 day/week for 4 weeks (subacute exposure), or a single 0.8 ppm exposure for 5 hrs (acute)
    - "Females were refractory to change." – the abstract of the paper says "Female rats appeared to be more affected than males to O-3 regardless of diet."
    - Actual results from the study:
        - Male Rats Acute: decrease in total cholesterol with ozone exposure in control group only; no change in glucose, HDL, or LDL; increase in triglycerides in control and high fat group but not in high fructose group with ozone exposure.
        - Male Rats Sub-acute: increase in HDL with ozone exposure in high fructose and high fat diets; no change in total cholesterol, glucose, LDL, or triglycerides.

- Female Rats Acute: decrease in LDL cholesterol with ozone exposure in control group and high fat diets; no change in glucose, HDL, triglycerides, or total cholesterol
- Female Rats Sub-acute: no change in total cholesterol, glucose, LDL, HDL, or triglycerides.

In section 5.2.3 (glucose and insulin homeostasis) evidence is presented from three studies. I have summarized the results from those studies below, having seen details that do not always agree with the EPA's summary:

- Bass et al. 2013:
  - Methods: 1, 4, 12, or 24 month old male Brown Norway rats were exposed to 0.25 or 1 ppm ozone for 6 h/day for 2 days (acute) or 6 h/day, 2 days/wk for 13 weeks (subchronic)
  - Acute exposure - all age groups (1, 4, 12, 24 mo) had increased glucose intolerance after 1 ppm ozone. Fasting blood glucose was elevated in 1, 12, 24 month old animals with 1 ppm ozone. 18 hrs after exposure glucose tolerance returned to baseline. No apparent effect of 0.25 ppm; increase in serum HDL at 1 ppm in 12 month-old rats only, no other changes in total or LDL cholesterol.
  - Subchronic exposure - all age groups had increased blood glucose after the glucose tolerance test compared to FA, but there was no increase in fasting blood glucose. No stat sig effect of 0.25 ppm. Increase in serum HDL at 1 ppm in 12 month-old rats only, no other changes in total or LDL cholesterol. Significant decrease in serum triglycerides in 0.25 but not 1 ppm exposure in 24 month old animals only.
- Miller et al. 2016b:
  - Male Wistar-Kyoto rats exposed to 0, 0.25, or 1 ppm ozone, 5 hrs per day, 3 consecutive days per week for 13 weeks.
  - Acute (1 week exposure) - fasting hyperglycemia and glucose intolerance at 1 ppm - reduced after 3 days of exposure compared to 1 day of exposure. No indication of impairment of peripheral insulin-mediated glucose clearance with ozone exposure. With 1 ppm ozone, no increase in insulin with glucose injection.
  - Subchronic - fasting hyperglycemia and glucose intolerance at 1 ppm - seems to be worse than after acute exposure. Completely reversed after 1 week of recovery. No indication of impairment of peripheral insulin-mediated glucose clearance with ozone exposure. With 1 ppm ozone, no increase in insulin with glucose injection. No change in liver or muscle insulin resistance.
- Gordon et al. 2013:
  - Male B/N rats exposed for 17 weeks, once per week 6 hrs per day for 17 weeks to 0 or 0.8 ppm ozone.

    - Serum insulin was significantly higher in old rats exposed to ozone, but not in 4-month old rats exposed to ozone. This is in contrast to the potential decrease in insulin seen in the Miller 2016 study. Ozone had no overall effect on body weight.
- Miller, Gordon, and Bass all came from essentially the same set of authors. But they show different effects: Bass showed no change in fasting glucose with subchronic exposure, but Miller did; Miller showed decreased baseline insulin in subchronic exposed adult animals, but Gordon showed no change in adult exposed animals, and increases in insulin in senescent exposed animals. The EPA should speak to whether there are patterns in these results, or if the differences are spurious or show strain differences.
- In section 5.2.4 (adiposity, weight gain, and obesity), the EPA states that "Serum lipids (triglycerides and HDL cholesterol) were significantly changed with ozone exposure in aged animals [24-month-old males, 0.25 ppm ozone, 6 hours/day, 2 days/week for 13 weeks; Bass et al. (2013)], an effect not seen in younger animals with the same exposure."
- However, in the Bass 2013 study they present that: subchronic exposure to 1 ppm was associated with increased HDL cholesterol in 12 month old animals only (not 1, 4, 24 months) or at 0.25 ppm, and there were no other cholesterol effects. Table 1 of the study shows a decrease in serum triglycerides in 24 mo animals with 0.25 ppm ozone but not 1 ppm, making this a suspect result (and this result not discussed by the study authors in the text).
- In section 5.2.5 (metabolic syndrome and type II diabetes), the EPA states for the Jerret et al. 2017 study that "The study observed increased hazard ratios for incident diabetes (1.28; 95% CI: 1.06, 1.55); however, when adjusted for NO2, this relationship was slightly weaker and had wider confidence intervals (1.20; 95% CI: 0.96, 1.50)." As per the study, these numbers should be 1.18 (1.04, 1.34) (basic model + 10% criteria), and adding NO2 changes the estimate to 1.13 (0.97, 1.31). The estimates provided in section 5.2.9 discussing copollutant confounding, and the results presented in Table 5-12 are also incorrect. The HAWC exposure summary is also incorrect, noting for exposure that "The 2-year average ozone concentration prior to the diagnosis or last follow up was assigned to each study participant.". However, only the 2007-2008 concentrations of ozone from the CMAQ model were used (although the follow up period was 1995-2011).
- In section 6.1.8 (causality determination), the EPA states that "Recent studies continue to assess the influence of potential confounders on the ozone-mortality relationship including copollutants, temporal/seasonal trends, and weather covariates; overall, these studies report that associations remain relatively unchanged across the different approaches used to control for each confounder." There is actually quite a bit of variability in responses from different models with copollutants and temporal/seasonal trends (particularly the latter), which is not adequately summarized by saying that they are "relatively unchanged". Relative to what?

Dose Concordance

- To consider a rough estimate of the effect of exercise on inhaled dose, I multiplied the ozone concentration by the fold-increase in exercise ventilation for the amount of time that the subjects were engaged in exercise.
    - Controlled human exposure studies have shown significant respiratory symptoms in healthy humans exposed to 72 ppb for 6.6 hours with exercise (5 hrs at 40 L/min + 1.6 hrs at ~10 L/min, at 0.072 ppm, ~**1.6 ppm·hrs**)
    - Epidemiology study of asthma hospital admissions in Texas: statistically significant increase in asthma hospital admissions in 5-14 year-olds associated with 1 day (lag 0) 8-hr maximum ozone concentrations (average 8-hr max concentration of 41.8 ppb). Using some mean values, children spend about 10% of their time outdoors; personal ozone concentrations are ~20% of outdoor ozone concentrations. Children in these age groups will on average spend ~7.5 hours at a ventilation rate of 11.35 L/min (about double the at-rest ventilation rate; Ref). So the average exposure is (1 hr at 2X rest ventilation rate at 0.042 ppm+7 hrs at 2Xrest at 0.0084 ppm; ~**0.2 ppm·hrs**). Even considering a 2-fold difference in BSA, there is a lot of dose difference amongst these studies.
- In section 3.2.3 (biological plausibility for long-term respiratory effects), the EPA does not provide the concentrations at which effects were observed. It is difficult to determine what effects are biologically plausible for occurrence in humans exposed to ambient concentrations of ozone (a typical annual average 8-hr max ozone concentration being in the range of 25-50 ppb), when experimental study results are presented without any concentration information. Concentration can substantially impact the effects of ozone.

Clarity of Presentation

- The EPA should note when discussing the results of controlled human exposure (CHE studies) done at 60 and 70 ppb, what the actual measured exposures were, not just the targeted exposures.
- Integrated Synthesis: "Under the assumption that respiratory symptoms might accompany similar ozone-induced FEV1 decrements, regardless of exposure duration, the model indicates that an 8-hour exposure to 64 ppb ozone concentration might reasonably be expected to cause an adverse response in young healthy adults." This statement mis-states "regardless of exposure duration", because it is then immediately followed by stating for "an 8-hour exposure". This should also include a statement that it assumes moderate to vigorous exertion for a substantial period of time.
- On page 3-26 the EPA describes pathways that researchers have demonstrated are activated in response to ozone to cause airway responsiveness. Are these pathways occurring simultaneously, or have different researchers found different pathways to be activated under different circumstances? More information should be provided here. Also

on page 3-26 the EPA notes that exposures of 0.8-2 ppm ozone cause airway hyperresponsiveness, but the time profiles of these should be included as well.

- In section 3.1.5.3 when the EPA discusses CHE studies, the EPA notes that in Horstman et al. 1995 they show that there are more wheeze symptoms in people with asthma exposed to ozone - is this compared to people without asthma, or ozone just increases wheeze in general? Also, the EPA states that "These observations are not changed by recently available studies or those in subsequent assessments (U.S. EPA, 12 2013a, 2006)." Does that mean that there are no other studies that have looked at this, or other studies have confirmed it?
- In section 3.1.5.4 (lung function) the EPA states that "it was concluded that individuals with asthma were at least as sensitive to acute effects of ozone as healthy individuals." This is an unhelpful statement. I don't think that anyone is arguing that people with asthma will be less sensitive to ozone. Similarly the EPA states that "However, despite limited evidence demonstrating increased sensitivity to ozone in individuals with asthma compared to those without asthma, there is consistent evidence that asthmatic individuals experience lung function decrements in response to acute ozone exposures." Again, I don't think that there is any argument that ozone causes auction lung function decrements in people with asthma, because it also causes decrements in people without asthma.
- Airway hyper-responsiveness (section 3.1.5.5) is potentially a very important endpoint because of concerns of people with asthma being particularly sensitive to ozone. The EPA notes that there are no differences in CHE studies in people with and without asthma for airway responsiveness with ozone exposure. The EPA seems to be making the argument that because people with asthma tend to have a higher baseline airway responsiveness, then they have less buffer against a potentially adverse effect. This is a valid argument and the EPA should make it more clearly. This section would also benefit from the inclusion of exposure details when discussing the CHE results.
- In section 3.2.1 (introduction to long-term ozone exposure and respiratory effects) the EPA presents the evidence from the last review as being very strong, and therefore does not make it clear why the causality determination was *likely* and not *causal*. This is the danger of only presenting positive results. Some explanation of the likely causal choice should be provided.
- Section 3.2.4.5 addresses severity of respiratory disease. However, the evidence seems to be the same as just general associations with respiratory disease. For example, data is presented for HAs and ED visits, and for association with symptoms – this doesn't provide information about increased severity with ozone concentration. I suggest that the EPA further clarify why this data specifically provides severity information, rather than dividing the data up as is usually done (i.e. a section on symptoms, a section on ED and HA visits, etc). In addition, it is noted in the causality determination section that these endpoints are typically considered short-term and the studies do not usually control for short-term exposure effects. Given these concerns the EPA should be clear about how to interpret these study findings.

- In section 4.1.7.4 the EPA notes that "Importantly, this study also found that after 18-hours ozone exposure, both levels of ozone reduced increased sensitivity to aconitine-induced arrhythmia ($p < 0.05$). Similar results were also found in an another study by this group (Farraj et al., 2016)." This statement is very confusing and seems to suggest that ozone decreases the sensitivity of rats to arrythmia induced by aconitine. However, the Farraj et al. 2012 study states that "both low- and high-O3 exposure significantly reduced the total dose of aconitine necessary to elicit the first ventricular premature beat relative to air-exposed controls." This makes it clear that ozone increases sensitivity to aconitine-induced arrythmia. The EPA should clarify this statement.
- In section 4.1.16 (effect modification) the EPA states in their discussion of life stage effects that "Cakmak et al. (2014) used a population of 8,662 Ottawa and Gatineau patients referred for 24-hour ambulatory cardiac monitoring with exposure linked to the monitor closest to their home address. In subjects over the age of 50 (n = 6,009) cardiac rhythm was disrupted by an increased presence of heart block (1.13; 95% CI: 1.01, 1.27)." It is not clear what is meant by "cardiac rhythm was disrupted by an increased presence of heart block", and what that means in the context of ozone.
- In section 5.1.2 (biological plausibility) the EPA notes that "Diabetes and metabolic syndrome are disorders of the autonomic nervous system." This statement should be explained and referenced.
- In section 5.1.4 (overweight and obesity) the EPA states that "Genetically obese mice had airway hyper-responsiveness and responded more vigorously to acute ozone exposure than did lean controls (Shore, 2007)." Responded more vigorously in what way? To AHR, or some other response?
- In section 4.1.9 (heart rate and heart rate variability) the EPA notes that heart rate is a key prognostic indicator. They should state what heart rate is a key prognostic indicator *for*.
- In section 5.1.4 (overweight and obesity) in the summary of the Gordon 2017a study, the EPA notes that the purpose of the study was to "To determine the role of maternal exercise and diet on ozone's effect on glucose homeostasis and obesity in offspring, a study was conducted with multiple diet and exercise options." Then they note that while ozone exposure increased the glucose intolerance at one or two time points in every group compared to their matched FA controls, "comparisons were not made across groups." Since the purpose of the study was purportedly to determine how maternal exercise and diet affects ozone's effect on glucose homeostasis, it doesn't make sense that the study authors didn't test the difference between groups.
- In section 5.1.5.1 (Other effects, inflammation) it isn't clear whether the peripheral inflammation is only present in animals that are obese or have diabetes, or if this is generally associated with ozone exposure. This should be clarified.
- At the beginning of section 5.1.5.3 (Other effects, endocrine hormones), conclusions are presented about the effects of ozone on the endocrine system with citations to the review article by Snow et al. (2018). However, this section should be presenting the evidence for the presence or absence of these effects, not citing conclusions from a review article

before presenting any data. I suggest moving this information to the end of the document, or removing it.

- In section 5.1.5.3 (Other effects, endocrine hormones), the EPA summarizes that "Elevated circulating stress hormones are consistently observed in animal models and in controlled human exposure studies after short-term ozone exposure." This should more precisely state that it has been observed in ONE CHE study.
- In section 5.1.5.4 (Other effects, serum lipids), the EPA states in reference to the Chen et al. (2016a) study, "No outcomes were reported for any metabolic endpoints evaluated with short-term increases of ozone exposure." If no outcomes were reported with short-term ozone exposure in the Chen 2016 study, then why is it covered here? Or does EPA mean that no metabolic outcomes were associated with short-term ozone exposure? If that is what EPA means, then they should say so.
- In section 5.2.2 (biological plausibility) there is a repeated sentence: "Ozone acts as a pulmonary irritant and stimulates nasopharyngeal and pulmonary nerves and receptors, including the trigeminal and vagal nerves, which induces downstream effects to the autonomic nervous system."
- In the first paragraph of section 5.2.11 (summary and causality determination) the EPA twice makes the statement "In prospective cohort studies in the U.S. and Europe increased incidence of type 2 diabetes was observed with long-term exposure to ozone."
- In section 6.2.3 (total non-accidental mortality) the EPA states "Generally, epidemiologic studies of long-term ozone exposure and total mortality use the 8-hour daily max ozone metric, though there are instances in some that use the 24-hour avg [e.g., Sesé et al. (2017)], or the 1-hour daily max [e.g., Jerrett et al. (2009)]." Presumably the EPA means here some long-term average of those different metrics is used in those studies?
- In section 6.2.4.1 (effect modification, pre-existing disease), the EPA seems to consider studies that do not compare people with the pre-existing disease, to people without the pre-existing disease (e.g. Zanobetti & Schwartz 2011 does not investigate people without a recent hospital admission, so they cannot conclude that people with a recent hospital admission have a higher risk than those without). The EPA should consider studies with the appropriate comparators to differentiate whether a pre-existing disease makes a person more sensitive to ozone exposure.
- Section 7.1.3 (pregnancy and birth outcomes) should present effect estimates and confidence intervals with epidemiology study results, as is done in other parts of the document.
- Section 7.2 (nervous system effects) should present ozone concentrations used in the animal toxicology studies in the text, as is done in other parts of the document for these types of studies.
- In section 7.2.2.5 (neurodevelopmental effects), the EPA notes that "…in the current ISA these data are only reviewed in the long-term exposure section due to the sensitivity of the developing nervous system to toxicants and the potential for long-term impacts." This doesn't make sense - the divisions of short and long-term are made based on the exposure

time, not the outcome. Mortality associated with short-term exposure could also be considered to have long-term impacts, but it is still included with short-term exposure. I agree with grouping together the neurodevelopmental effects, and I think that the long-term category is fine because in humans this can be up to a 9-month (i.e. not acute) exposure, but not because it may have long-term consequences.

Consistency of Results

- In section 3.1.5.2 the EPA notes that for asthma ED visits, that the evidence is consistent for both children and all ages - this is not what is typically seen for HA - is there a reason for this discrepancy?
- In section 3.1.7.2 the EPA notes that data from CHE studies suggest that ozone exposure activates the adaptive immune system, which may bolster defenses against infection. However, the epidemiology study results indicate an association between ozone and increased respiratory infections. The EPA should provide some discussion of how they reconcile these results.
- In section 5.1.5.4 (Other effects, serum lipids), the EPA summarizes that "Ozone exposure alters serum cholesterol in multiple animal models." However, the magnitude and direction of those alterations is not consistent. Some studies show increased HDL, some LDL, some no changes, and some CVD models show decreases in HDL or LDL. The EPA should discuss what these differential results mean in terms of the ozone mode of action and adversity of effects.
- In section 5.2.4 (adiposity, weight gain, and obesity), the EPA concludes from the epidemiology study results that "Recent evidence is limited, but provides some evidence that long-term exposure to ozone is associated with increased weight gain and obesity" However, the highest quality, most relevant to the US study showed no effect. In addition, the two Chinese studies that did potentially show effects did not have a study quality evaluation conducted by EPA - should the general conclusion be based on those studies?

Applicability of Animal Study Results

- The animal studies being cite in section 3.2.4.1 (new onset asthma) are being conducted at concentrations WAY HIGHER than ambient, and do not seem to be investigating a no effect level, which makes it difficult to know how comparable these effects are in humans. And even at these high concentrations effects weren't always consistent, or showed a dose response (e.g. Chou 2011 did not show increased BALF eosinophils with 5 biweekly 8-hr exposures to 500 ppb ozone, but Crowley 2017 did observe BALF eosinophils with 11 biweekly exposures). Lee et al. 2011 demonstrated evidence of a threshold in rat studies of post-natal exposure, with less frequent exposures to 500 ppb

ozone for 6 hrs per day having no developmental effects, compared to more frequent exposures.

- For the lung function and development section 3.2.4.2 the EPA presents animal studies of high exposures over multiple weeks causing lung function developmental effects. Seriously, of course chronic exposure to that high a concentration of ozone is going to affect lung development, we have already established that it can cause lung damage. It would be far more useful if a no and low-effect concentration was established. The rodent models do provide some useful information that there is a low concentration below which effects were not observed. The EPA should provide information about how this informs possible effects in human children?
- In section 4.1.4 (heart failure, impaired heart function) the EPA notes in the animal toxicological section that there were effects observed with ozone exposure in isolated, perfused hearts (McKintosh-Kastrinksy 2013). If this study involved direct exposure of the heart to ozone, how is this related to CVD effects that occur with pulmonary ozone exposure? The EPA should provide more information about how this type of study is relevant and how to interpret the exposure dose.
- In section 5.1.2 (biological plausibility) the EPA discusses the hyperthermic/hypothermic response in adult male rats caused by ozone. The EPA should provide more information about this to demonstrate its relevance to humans. For example, is this response also seen in other species/sexes? Is this thought to be comparable to humans? Is it dose-responsive?
- In section 5.1.2 (biological plausibility) the EPA notes the changes in thyroid hormone with ozone exposure in rodents. Given the known differences between circulating thyroid hormone and thyroid hormone responses on rodents compared to people, is there evidence that this mechanism would be relevant in humans?

<u>Interpretation of Study Results</u>

- In section 3.2.4.2 (lung function and development) the EPA notes that "To characterize lung health, lung function metrics capture the cumulative effects of pulmonary growth, damage, and repair (Wang et al., 1993). As such, measures of lung function are effective indicators of pulmonary effects related to exposure to environmental stressors." How are effects on lung function of acute exposures separated from effects of chronic exposures in these studies? That is, how do the authors know that they are measuring a long-term effect over time, and not just capturing the effects of the last day's exposure? The EPA should clarify these interpretations.
- The CVD plausibility section is supported by activation of the sympathetic nervous system, but in the respiratory section activation of the parasympathetic nervous system is cited. Can these occur simultaneously?
- In section 5.1.3 (glucose and insulin homeostasis) the EPA states that "Recent epidemiologic studies provide some evidence of associations between short-term ozone exposure and these endpoints (Table 5-5)." However, the evidence presented is mostly

null, or in opposite directions (one study showing increasing fasting glucose, another showing decreasing fasting glucose). Also, what is the significance of the presented changes in fasting glucose (e.g. 0.5-3% changes)? It seems that the normal range is something like 65-90, which is an up to 75% variability. It also seems that these factors would be highly affected by diet. Is diet controlled for in any way in these epi studies?

- In section 5.1.5.3 (Other effects, endocrine hormones), the EPA summarizes "In healthy rodent models, short-term ozone exposure was associated with either elevated serum leptin (Miller et al., 2015; Bass et al., 2013; Sun et al., 2013) or a trend toward increased leptin (Gordon et al., 2017b). In obese animals (Zhong et al., 2016) and diabetic animals (Ying et al., 2016), there was significantly decreased serum leptin with ozone exposure. Thus, healthy and diseased animal models have significantly different leptin responses to ozone exposure or the temporality differences between studies might explain the directionality differences." What is the potential significance of these differences in leptin between health and obese/diabetic animals, and how can they be interpreted?
- In section 6.1.5.4 (effect modification, temperature), Figure 6-4 shows that there were some very high hourly ozone concentrations in this dataset (consistent with the data coming from 1987-2000) - are these patterns still relevant and observed at lower current ambient concentrations?
- In section 5.1.5.2 (Other effects, liver effects), the EPA presents metabolomic and gene and protein effects from several studies (Miller et al. 2015 and Theis et al. 2014). Were the results from these studies consistent with one another in terms of the pathways affected by ozone?
- In section 5.1.4 (overweight and obesity) the EPA states that "Studies done at the U.S. EPA examining effects of ozone at various concentrations (0.25, 0.5, 1.0 ppm) in healthy and obese rat models with leptin receptor mutation and associated cardiovascular disease demonstrated low sensitivity to ozone-induced lung injury and neutrophilic inflammation (Kodavanti, 2015)." I think that this is the wrong reference. Also, the statement suggests that animals with the leptin receptor mutation were less sensitive to ozone-induced lung injury and inflammation? Does this mean that EPA thinks that leptin is involved in the ozone-induced damage? Is this hypothesis in the respiratory chapter biological plausibility section? The EPA should consider separating out the results from these different rat models to be clear about what effect was seen in which model.

<u>Completeness of Study Information</u>

- In section 5.1.4 (overweight and obesity) the EPA states that "Studies done at the U.S. EPA examining effects of ozone at various concentrations (0.25, 0.5, 1.0 ppm) in healthy and obese rat models with leptin receptor mutation and associated cardiovascular disease demonstrated low sensitivity to ozone-induced lung injury and neutrophilic inflammation (Kodavanti, 2015)." I think that this is the wrong reference.

- At the beginning of section 5.1.5.2 (Other effects, liver effects) the EPA states that there is evidence for changes in the gut microbiome. However, there doesn't seem to be any direct evidence presented for that effect. Is this conclusion based on the change in bile acids? If so, that should be stated with appropriate caveats about the interpretation of this evidence, or the conclusion should be removed unless direct evidence is provided.
- Appendix 5 requires copy editing as well as careful attention to detail on exposure and experimental details.

Causality Determinations

For the epidemiology studies cited in this ISA, it seems that there is more work to be done on separating the effects of the many variables that can contribute to the measured endpoints. My question to the expert panel about the appropriate control for confounding, colliding, and mediating questions generated a general assessment that this kind of careful variable control maybe is not used frequently and remains an uncertainty in interpreting epidemiology study results.

Long-Term Ozone Exposure and Metabolic Effects (5.2). To justify the conclusion of "likely to be causal", the EPA states (in section 5.2.11):

- "Consistent epidemiologic evidence of increased risk of diabetes or metabolic syndrome" - based on Jerret et al. 2017 results from the Black Women's Health Study Cohort; Renzi et al. 2017 results from the Rome Longitudinal Study Cohort; and Yang et al. 2018 from the 33 Communities Chinese Health Study Cohort.
    - Only one of these studies was assessed for study quality by the EPA (Jerrett) and it had incorrect study details in the ISA and the HAWC study quality assessment. That study showed a loss of significant association between ozone concentrations and incidence of type II diabetes when NO2 was included as a copollutant. Renzi et al. shows marginal associations with ozone at best (1.01 CI 1.00 to 1.02), and only in people under 50, but not older than 50. Yang is a Chinese study with substantial copollutant and other confounding concerns as noted by the EPA (these caveats were not included in section 5.2.11).
    - The EPA also cites results from Malmqvist et al. 2015 of effects on gestational or type I diabetes from developmental exposure to ozone (from section 7.1.3). However, this endpoint is not discussed in the causality designation, the study was not discussed in section 7.1.3, there was no statistically significant effect of ozone on gestational or type I diabetes, and there was no study quality evaluation.
- The EPA notes that the aforementioned findings are "consistent with two long-term ozone exposure studies in China, one in adults and one in children, presented increased odds of obesity in both adults and children as obesity is a risk factor for type 2 diabetes (T2D)."

    - These results are not consistent with results from the Black Women's Health Study Cohort (the same cohort as Jerret et al. 2017) that did not find an association between ozone and body weight gain (White et al. 2016). White et al. 2016 had a study quality evaluation in HAWC, while the two referenced Chinese studies (Dong et al. 2014, Li et al. 2015) did not.
- "Epidemiologic evidence of increased diabetes-associated mortality" – Based on diabetes mortality in the ACS cohort (Turner et al. 2016) and CanCHEC cohort (Crouse et al. 2016)
    - The EPA did not conduct a study quality evaluation of these studies. Why was that?
- There are a "limited number of epidemiology studies that evaluate potential copollutant confounding for PM or NOx" (and Jerrett et al. 2017 shows a loss of statistical significance of the effect when considering NOx confounding)
- "Animal toxicological studies of impaired glucose tolerance, fasting hyperglycemia, dyslipidemia, insulin resistance, and activation of the neuroendocrine pathway with ozone exposure"
    - The EPA summarizes the ozone concentrations associated with effects as 0.25 and 1 ppm, however, the available animal toxicology studies (Miller et al. 2016b, Bass et al. 2013, and Gordon et al. 2013) do not show effects at 0.25 ppm and Gordon et al. 2013 uses 0.8 ppm.
    - The animal evidence is not always summarized correctly, shows somewhat inconsistent results, and the 3 cited studies were conducted largely by the same group of authors.
- The EPA states that "The mortality findings are supported by epidemiologic and experimental studies reporting effects on glucose homeostasis and serum lipids, as well as other indicators of metabolic function (e.g., peripheral inflammation and neuroendocrine activation)." Studies of peripheral inflammation and neuroendocrine activation were not discussed in this chapter. The EPA also states that "In addition, these pathophysiological changes were often accompanied by increased inflammatory markers in peripheral tissues, and activation of the neuroendocrine system (Section 7.2.1.5)" However, the single study discussed in section 7.2.1.5 was a gene expression study after a single acute dose of ozone, with a return to baseline expression levels by 24-hours after exposure.
- The EPA states that "The animal toxicological studies provided evidence that long-term ozone exposure resulted in impaired insulin signaling, glucose intolerance, hyperglycemia, and insulin resistance (Section 5.2.3.1). In addition, these pathophysiological changes were often accompanied by increased inflammatory markers in peripheral tissues, and activation of the neuroendocrine system (Section 7.2.1.5)" This statement lacks all the nuances of these studies, and the inconsistencies, and the fact that there was no indication of impairment of peripheral insulin-mediated glucose clearance with ozone exposure (Miller et al. 2016b).

- The EPA states that "Importantly, short-term ozone exposure studies also provided evidence that ozone exposure could contribute to the development of metabolic syndrome and show consistency with the evidence that long-term ozone exposure could lead to development or worsening of metabolic syndrome or its risk factors." However, there are several studies that investigated acute and subchronic exposure, and it was not clear that there were worse metabolic responses with increased exposure time (Miller et al. 2016b, Gordon et al. 2016b).

In section 7.1.2.2 (male reproduction), the EPA presents just a few epidemiology studies with somewhat conflicting results, and a single animal toxicology study. However, their conclusion is "Overall, there is evidence of impaired spermatogenesis and decreased sperm count and concentration from epidemiologic studies, and decremental effects on testicular morphology and impaired spermatogenesis from toxicological studies with ozone exposures." This is overly certain based on the available database.

<u>Biological Plausibility</u>

In section 7.1.2.1 (male and female reproduction and fertility, biological plausibility), the EPA does not make clear how the pathway leads from systemic inflammation to testicular and pregnancy effects. There is not connection provided there, and the "how" of the connection should be noted.

In section 7.1.3.1 (pregnancy and birth outcomes, biological plausibility), the EPA provides very little data to support the biological plausibility pathway. Further, there is not convincing evidence of systemic inflammation and oxidative stress induced by ozone exposure (a necessary upstream event in the pathway). The only statement supporting this upstream event is "The initial event of altered systemic oxidative stress is demonstrated in the epidemiologic literature with ozone-dependent increased odds of elevated CRP levels in nonpregnant individuals but CRP was unchanged at GD 5 in ozone exposed pregnant rodents (Miller et al., 2019)." There is no reference for the epidemiology study, and it provides much weaker evidence than the animal study.

**<u>Questions for Consultants</u>**

*The EPA states in the ISA preamble that "Traditionally, statistical significance is used to a larger extent to evaluate the findings of controlled human exposure and animal toxicology studies. Understanding that statistical inferences may result in both false positives and false negatives, consideration is given to both trends in data and reproducibility of results. Thus, in drawing judgments regarding causality, the U.S. EPA emphasizes statistically significant findings from experimental studies, but does not limit its focus or consideration to statistically significant results in epidemiologic studies."*

***1) It has been established that associations found in an epidemiology study can be due to: causation, bias, chance, and/or confounding. If the concept of statistical significance is not useful in epidemiology studies, then how do the study authors/EPA rule out that chance has caused the observed association?***

Responses:

Dr. Jaffe: "Regarding statistical significance in the preamble, I think the line right above your quoted line in the pre-amble is also important in this context "Other indicators of reliability such as the consistency and coherence of a body of studies as well as other confirming data may be used to justify reliance on the results of a body of epidemiologic studies, even if results in individual studies lack statistical significance." To me, what this is saying that if a group of studies, each with significance at say a 90% confidence all demonstrate a similar effect, then this can be used to justify a conclusion, even if no single study reaches a 95% confidence."

Dr. Jansen: "I was not aware of this. Imposing "significance" as a criterion on one type of study and not others seems wrong. On the other hand, does it explain the use of the phrase "positive associations" in many places (e.g., see page 4-2) rather than "significant positive associations?" I find such phrasing troublesome as it could be implying more rigor than exists and lacks clarity. If the result is null, it's null. Significant results should be shown distinctly, not diluted by "positive" null results, and weighted more heavily."

Dr. Lipfert: "Essentially, they do not. I take strong issue with the above EPA protocol on statistical significance and note that studies with wide confidence intervals are included. Experimental studies involve defined exposures with no co-pollutants or temperature effects and no measurement error. Variations among studies are thus real and should be analyzed as such. By contrast, epidemiological studies are subject to all of these sources of bias, especially temperature and the treatment of lags. Foley et al. (2003) reported that EPA considered outdoor ambient air quality from fixed ambient monitoring sites to be a "surrogate for exposure", which requires consideration of indoor exposures, by contrast with experimental studies."

Dr. North: "Statistical significance is useful in epidemiological studies, but in a limited way. These studies use regression to determine the association between a predictor x and a consequence y. This association might be statistically significant, that is, a good predictor, but causality could be absent: There may be a cause z that affects both x and y. Two examples: Children's shoe side predicts the children's reading ability. (Example due to Judea Pearl, in The Book of Why). Ice cream sales predict heat stroke cases. (I think Dr. Cox uses this one, which was more accurate before the era of air conditioning.) Progress for better prediction is to consider that there may be other factors that are predictive of y, get the data on these, and use these data on making the prediction. Children's age and high ambient temperature are candidates for the two examples."

Dr. Parrish: "I can respond to this question as a scientist with extensive experience in interpretation of results based on statistical significance, but not as an epidemiological expert. The concept of statistical significance is useful in interpreting the results of most scientific studies, but it is not of yes-or-no utility. Recent literature (e.g., Amrhein et al., 2019; Hurlbert et al., 2019) emphasizes the importance of not using statistical significance in a dichotomous manner, e.g. to decide whether the results of an analysis rules in or out any particular cause of an observed association. It is recommended to simply present p-values without label or category. With regard to the present question, the results of any particular epidemiology study can provide an estimate of the probability that chance has caused the observed association, but that probability can never be reduced to zero. The same statement can be generally applied to the results of all studies that attempt to understand the cause of an observed correlation or association."

Dr. Thomas: "In addition to bias (of which confounding is one kind), chance can certainly lead to non-causal associations. Assessment of statistical significance is essential to judge the likelihood that an association could be due to chance, so it's incorrect to say that it "is not useful in epidemiology studies." Despite the longstanding and on-going debates about the usefulness specifically of p-values for this purpose (Greenland et al. 2016, Wasserstein and Lazar 2016), as opposed to a variety of other approaches (e.g., confidence intervals, Bayes Factors, etc.), they remain the most commonly used method for judging the possibility of chance. I do not see that the EPA has dismissed statistical significance testing in its evaluation of the evidence, although they correctly do incorporate "trends in data and reproducibility of results" as well as other considerations in their evaluation of the epidemiologic evidence."

Dr. Sax: "Epidemiological studies cannot be the sole basis for establishing a causal relationship because of the inherent limitations and because, in the case of observational epidemiological studies, you cannot rule out bias, chance and/or confounding with reasonable confidence. The issue is even more difficult when the observed effects are very small and not statistically significant or marginally significant. Because of these limitations it is essential to evaluate all lines of scientific evidence, including experimental evidence (human chamber studies, animal studies, mechanistic studies). By evaluating the consistency and coherence within and across the various scientific lines of evidence, one can obtain a better picture of whether a causal association is more or less likely. More importantly, the evidence may be able to also elucidate levels of exposure at which effects are more likely."

<u>My summary of the responses:</u> The general conclusion from the expert consultants was that statistical significance does need to be given some consideration, but there are also other factors such as patterns in the epidemiology data that also should be factored into the conclusions that are drawn.

*Some short-term epidemiology studies use a method that is termed "case-crossover". These studies assess the pollutant concentration on the day of a health effect, and "control" days are those days when a person did not experience that health effect. My understanding is that the intention of this method is to control for intra-individual confounders. These study designs often use days before and after the health event (often matched to day of the week) as control days.*

***2) Am I correct in understanding that the intention of ozone case-crossover studies is to compare the ozone concentrations on a day when a health effect occurred for a person, to the ozone concentrations on a day when that health effect did not occur for that person?***

***3) If so, then it would be important that some other factor (not related to ozone) did not prevent the health event from occurring on a control day. These studies often use days before and after the health event as control days, but for mortality studies (such as Di et al., 2017), how can a day after death be used as a control day? It doesn't matter what the ozone concentrations are after a person's death, that person would not be able to respond to that concentration. How should we interpret case-crossover studies that use control days after the event (particularly mortality) occurred?***

<u>Responses:</u>

Dr. Jansen: "2) I believe that is correct. I also seem to recall use of the ozone on, say, another Tuesday in a month, assuming the event happened on a Tuesday. 3) I believe your concern is not limited to mortality. I would expect using the day after for a hospital admission is affected by medical treatment and being confined indoors. If this is a fatal flaw, the studies should be excluded. If not, then there is clearly an uncertainty and the study results should be down weighted."

Dr. Lipfert: "2) Yes. 3) I am not an expert on case-control studies but I question their use in temporal rather than spatial studies. Ozone has both diurnal and seasonal trends as does ambient temperature, its primary confounder (see the Figures 2 and 3 above). Further, acute effects on mortality persist for several days, perhaps up to a week. Thus "case" days may not be independent of "control" days. Different people with different characteristics die on each day for different reasons. I would rather see both case and control periods extended for say, 3 days or more. The time-series model of Murray and colleagues (not cited in the ISA) considers temporal patterns of subject frailty, ambient temperature and air quality, each over several days. These temporal patterns are much easier to interpret than case-control findings. Di et al. combined spatial and temporal analyses, thus introducing geographic and climate variability (for no particular reason). I would have much rather seen a conventional time-series analysis in each of several locations involving socioeconomic differences as well."

Dr. North: "2) I'm not a fan of these cross-over studies, especially for mortality as the end point. For a good example, consider exposure to wildfires such as we have been experiencing in California. Consider hospital admissions for respiratory distress. What was the level the day before the smoke plume affected the area? What was the level the day the plume arrived? And after the wind blew the smoke away, then what was the level the next day? Yes, one might expect a low response level before the plume, and high levels after the high exposures, perhaps persisting for days after the levels have dropped. For Q3, the response is number of deaths on the exposed day versus the control day, and not the death of an individual person. (With cohort studies, it is more complex.) 3) I read the Di et al. study. Co-authors Schwartz and Zanobetti are among those trying to figure out how to do epidemiology where additional factors are considered. But I am not persuaded that confounding was not a significant issue for the results in the Di et al. study. The data base was all Medicare patients who died in a twelve year period. Most of the deaths occurred on days with ozone and PM2.5 levels well below the current standards. The death rates per 10 μg/m3 for PM2.5 was 1.45 per million persons at risk per day, and for 10 ppb ozone, 0.66 per million. These are extremely small numbers, but with sample size of nearly a hundred million days, the confidence limits were narrow around these numbers and did not include no increased risk. I looked up reference 9, Maclure (1991), on the study design. The Maclure Abstract begins, "A case-control design involving only cases may be used when brief exposure causes a transient change in risk of a rare acute-onset disease." I don't see the biological plausibility of comparing case days and control days for total mortality – not a rare acute–onset disease, but rather a situation where people who may be already very sick tend to die on days when they have additional stress. I suspect that high temperature may have acted as a confounding variable. Looking at figure 5 in the Di paper, I notice that the exposure response curve seems to flatten out (i.e., is insensitive to exposure level) for the higher 50% of the exposures, both for ozone and for PM2.5 . If these pollutants were causing the mortality increase, I would expect that the lower half of the exposure levels would be the flatter portion, and at higher exposures there would be more of a positive concentration response relationship. What may be going on is that in the days with higher half of the exposure levels, the pollutant levels are correlated (but rather poorly) with the frequency of very hot days. On such very hot days mortality is significantly elevated. But on the lower pollution half of the days, there is a stronger correlation: a much lower frequency of very hot days. Very hot days can cause stress to an elderly person, especially in non-air conditioned space. Remember, the Di et al response rates are on the order of a one-in-a-million change. A small number of very hot days correlated with elevated exposure levels might give results such as reported in this paper."

Dr. Thomas: "2) Yes, that is the correct interpretation. An advantage of this design is that by making comparisons with an individual, between-individual confounding is completely eliminated, as are any factors that do not vary over time. While factors other than pollution that do vary over time, like weather, could still be confounders, these can be controlled in the analysis by standard statistical adjustment methods, as in case-control or time-series studies. 3) The original case-control design (Maclure 1991) involved a comparison of exposure at the time

of the event (or some pre-specified time prior to it to allow for lag effects) to that at some previous comparable ("referent") time. For example, the referent time might involve the same day of the week to control for systematic weekly variation in pollution levels and/or confounders. My colleague, Bill Navidi (1998) pointed out, however, that seasonal variation and especially long-term trends in pollution levels could lead to bias if referent times always preceded event times, even if one or more entire year cycles were included; while there would be no bias if there were no long-term trends and if pollution followed a perfectly symmetric (e.g., sinusoidal) seasonal pattern, departures from such symmetry, as are common for both pollution and meteorology, would lead to bias. Instead, he proposed the "bidirectional case-crossover" design, in which two referent times, one before and one after, equally spaced around the event time, are used. The original Macluer design was intended to study personal time-varying characteristics such as behaviors that could be "triggers" for an event like death or heart attack; in this setting, it would be impossible to observe a behavior that occurred after death! In air pollution studies, however, personal behaviors are not being studied, but ambient exposures are and these can be measured and used meaningfully for comparison. While it is obviously true that pollution after the event could not be causally related to the event, the purpose of this design is to get an unbiased estimate of the expected exposure at the time of the event for comparison with the actual exposure at that time and can be interpreted as a sampling-based analog of the standard time-series approach for acute effects (Bateson and Schwartz 1999, Fung et al. 2003, Lu and Zeger 2007). Various versions of this design have subsequently been widely adopted in air pollution studies. Although the original bidirectional design has subsequently been shown to be slightly biased (Lumley and Levy 2000), a modified version involving using fixed time-strata, comparing exposures at event times within each stratum with those at all or selected times (e.g., day-of-week matched times) within the same stratum before and after the event, has been shown to be unbiased (Levy et al. 2001a, Janes et al. 2005a, Janes et al. 2005b), and this design has become the standard in substantive studies (e.g., (Levy et al. 2001b, Di et al. 2017)). As Mittleman (2005) says, "this strategy should be considered the de facto standard approach to the analysis of data arising in studies of the short-term effect of air pollution and weather" (see also references therein for additional studies using this design)."

Dr. Sax: "2) I believe that is the correct interpretation. 3) I think that this is a valid question, as you described by selecting days post-health effect this would violate an important epidemiological tenant for assessing a causal relationship – that is, that the exposure must precede the effect."

***4) What is the importance of dose-concordance in establishing the biological likelihood of ozone-mediated effects occurring at relevant exposure concentrations in humans? Particularly in the context of known dose information about ozone: total inhaled dose includes concentration, exposure time, and exercise duration; Hatch et al., (2013) have shown that humans and rats that are exposed to ozone at rest achieve similar alveolar ozone doses, and that humans exercising at 5-times a***

***resting ventilation rate achieved an ~ 5-times higher alveolar ozone dose; and that ozone concentrations are 2-10 times lower indoors where people spend most of their time.***

Responses:

Dr. Lipfert: "I fail to see any relevance. Controlled animal and human clinical studies serve only to show what might happen under controlled and idealized conditions. Null findings may thus be the most important. By contrast, epidemiology shows what actually does happen in the real world, including variability."

Dr. North: "That is a good question. My impression is that at ppm levels, humans and rodents are about equally sensitive, and that prolonged exposures at 5 ppm or higher are life-threatening to humans. There seems to be some information at lower levels. See my responses to Drs. Cox and Packham."

Dr. Thomas: "If by "dose-concordance" you mean comparability of doses to animals and humans from similar external concentrations and ventilation rates, I would expect that there are so many factors that differ that it would be unreasonable to expect the same dose-response relationships, even if doses could be scaled in comparable units."

Dr. Sax: "Particularly in the context of known dose information about ozone: total inhaled dose includes concentration, exposure time, and exercise duration; Hatch et al., (2013) have shown that humans and rats that are exposed to ozone at rest achieve similar alveolar ozone doses, and that humans exercising at 5-times a resting ventilation rate achieved an ~ 5-times higher alveolar ozone dose; and that ozone concentrations are 2-10 times lower indoors where people spend most of their time. I think this is a very important issue and one that has not been resolved or evaluated by EPA in weighing the evidence across different studies (i.e., human chamber studies and animal studies)."

***5) Is there evidence that the animal models used to assess ozone effects (largely rats, mice, and non-human primates) are more, less, or similarly sensitive to ozone-mediated adverse effects compared to humans, at approximately equal inhaled doses?***

Responses:

Dr. Lipfert: "I don't know but, in my view, such tests should be only qualitative and used to study mechanisms. I don't see them useful to look for a "safe" dose since the actual human doses used in epidemiology remain unknown in part because of indoor effects."

Dr. Sax: "Again, this is a valid question, and given the evidence as presented in the ISA it is difficult to answer. I think that EPA's assertion that some of the high exposure levels used in the animal studies (based on the Hatch et al., 2013 study) are relevant to ambient exposures in humans is likely to be simplistic at best, and a more detailed analysis to support an answer to this question is warranted."

My summary of the responses: Given that the causality determination for metabolic effects of ozone exposure is mostly derived from animal toxicological studies, it is appropriate for the EPA to more thoroughly discuss the dosimetric similarities and differences between animals and humans, beyond the simplistic reference to Hatch et al. 1994.

*In this ISA I did not find population studies that considered causal pathways when assessing the association between ozone and health endpoints. It has been shown that the type of interaction between variables (e.g. confounding, colliding, mediating) can impact the results of regression analyses if these variables are controlled for in the regression equation.*

***6) In the absence of a causality diagram to direct the choice of variables to control in an epidemiological study, how can we judge whether a study has appropriately controlled for confounders, and has not inappropriately controlled for colliders (which can open up pathways between variables that otherwise would not be connected) or mediators (and thereby controlled away the effect)?***

Responses:

Dr. Jansen: "Clearly, the issue should be a key criteria used in the selection of and evaluation of studies. As I stated above, EPA needs clear criteria for study inclusion/exclusion, study quality, and causality classification. It is also not always clear which evidence is being given more weight than other evidence."

Dr. Lipfert: "I see causality and confounders in simpler terms: causality in terms of experiments and physiology and confounders in terms of bi-variate correlations and model evaluations with and without potential confounders. As stated above I believe that most of the mortality relationships are short-term and thus with few potential confounders. I don't think one can ever be sure that all of the long-term potential (spatial) confounders have been adequately controlled."

Dr. North: "Causality diagrams are still rare. But some epidemiology studies do consider multiple predictive factors, and explain how they do it. I expect we will consider this aspect in the risk assessment in the upcoming Ozone PA."

Dr. Thomas: "Very good question! Directed Acyclic Graphs (DAGs) can be useful tools for visualizing hypothetical relationships among observed and latent variables and for structuring an

appropriate analysis strategy (Greenland et al. 1999). Investigators typically have such pictures in mind when conducting an analysis, although they are seldom presented formally in a substance matter publication (they are more commonly included in statistical methods papers). The basic principles that confounders must be controlled using the best available data on known risk factors (or surrogates for unmeasured factors in an attempt to minimize residual confounding), and that intermediate variables on a causal pathway from exposure to disease not be adjusted for, nor for colliders (that are determined by both exposure and disease but are not causal for disease), are well understood. The art is in deciding which variables are or are not appropriate to adjust for. While there are a variety of formal statistical methods for dealing with adjustment uncertainty (Maldonado and Greenland 1993, Greenland 1996, Viallefont et al. 2001, Crainiceanu et al. 2007, Pope and Burnett 2007), it remains a matter for expert judgment, both by the original investigators and by critical readers.

Dr. Sax: "I don't think this is a new issue and as noted above, this is a particularly important limitation of observational air pollution studies. The study summaries that EPA presents in the ISA fall short of identifying the various limitations in the epidemiological literature and this remains an area of weakness in the overall evaluation of ozone health effects. For determining plausible (but not necessarily absolute) causation, a full integration of all lines of evidence is necessary. As noted previously, relying on only epidemiological evidence is not sufficient."

My summary of the responses: For the epidemiology studies cited in this ISA, it seems that there is more work to be done on separating the effects of the many variables that can contribute to the measured endpoints. My question to the expert panel about the appropriate control for confounding, colliding, and mediating questions generated a general assessment that this kind of careful variable control maybe is not used frequently and remains an uncertainty in interpreting epidemiology study results.

## References

Anderson, G.B., Bell, M.L., 2010. Does one size fit all? The suitability of standard ozone exposure metric conversion ratios and implications for epidemiology. J. Expo. Sci. Environ. Epidemiol. 20, 2–11. https://doi.org/10.1038/jes.2008.69

Brauer, M., Brumm, J., Vedal, S., Petkau, A.J., 2002. Exposure misclassification and threshold concentrations in time series analyses of air pollution health effects. Risk Anal. 22, 1183–93.

Cox, L.A. (Tony), 2018. Effects of exposure estimation errors on estimated exposure-response relations for PM2.5. Environ. Res. 164, 636–646. https://doi.org/10.1016/j.envres.2018.03.038

Folinsbee, L.J., McDonnell, W.F., Horstman, D.H., 1988. Pulmonary function and symptom responses after 6.6-hour exposure to 0.12 ppm ozone with moderate exercise. J. Air Pollut. Control Assoc. 38, 28–35. https://doi.org/10.1080/08940630.1988.10466349

Hatch, G.E., Mckee, J., Brown, J., Mcdonnell, W., Seal, E., Soukup, J., Slade, R., Crissman, K., Devlin, R., 2013. Biomarkers of dose and effect of inhaled ozone in resting versus exercising human subjects: Comparison with resting rats. Biomark. Insights 8, 53–67. https://doi.org/10.4137/BMI.S11102

Lange, S.S., 2018. Comparing apples to oranges: Interpreting ozone concentrations from observational studies in the context of the United States ozone regulatory standard. Sci. Total Environ. 644, 1547–1556. https://doi.org/10.1016/j.scitotenv.2018.06.372

Lipfert, F.W., Wyzga, R., 1996. The effects of exposure error on environmental epidemiology., in: Proceedings of the Second Colloquium on Particulate Air Pollution and Human Mortality and Morbidity. pp. 295–302.

McCant, D., Lange, S., Haney, J., Honeycutt, M., 2017. The perpetuation of the misconception that rats receive a 3–5 times lower lung tissue dose than humans at the same ozone concentration. Inhal. Toxicol. 29. https://doi.org/10.1080/08958378.2017.1323982

McDonnell, W.F., Kehrl, H.R., Abdul-Salaam, S., Ives, P.J., Folinsbee, L.J., Devlin, R.B., O'Neil, J.J., Horstman, D.H., 1991. Respiratory response of humans exposed to low levels of ozone for 6.6 hours. Arch. Environ. Health 46, 145–50. https://doi.org/10.1080/00039896.1991.9937441

Rhomberg, L.R., Chandalia, J.K., Long, C.M., Goodman, J.E., 2011. Measurement error in environmental epidemiology and the shape of exposure-response curves. Crit. Rev. Toxicol. 41, 651–671. https://doi.org/10.3109/10408444.2011.563420

Watt, M., Godden, D., Cherrie, J., Seaton, A., 1995. Individual exposure to particulate air pollution and its relevance to thresholds for health effects: a study of traffic wardens. Occup. Environ. Med. 52, 790–2.

Yoshimura, I., 1990. The effect of measurement error on the dose-response curve. Environ. Health Perspect. 87, 173–8.

Zaccai, J.H., 2004. How to assess epidemiological studies. Postgrad. Med. J. 80, 140–147. https://doi.org/10.1136/pgmj.2003.012633

# Dr. Ronald J. Kendall

The "External Review Draft" integrated science assessment for ozone and related photochemical oxidants (hereafter referred to as "draft ISA") prepared by the United States Environmental Protection Agency's (U.S. EPA) National Center for Environmental Assessment – Research Triangle Park Division (NCEA – RTP) as part of EPA's ongoing review of the primary (health-based) and secondary (welfare-based) National Ambient Air Quality Standards (NAAQS) for Ozone was released on September 26, 2019. The EPA staff were directed by EPA Administrator Andrew Wheeler to "continue progress on the review of the NAAQS for ground level ozone through production of the Draft Ozone ISA and accelerating the development of a Draft Ozone Policy Assessment so that both documents could be delivered for CASAC and public review by October 2019. The present preliminary comments by Dr. Ronald J. Kendall will firstly address Appendix 8 of the draft ISA for ozone and secondly will address Appendix 9.

To supplement the standardized charge question and guide the scientific review of this ISA, the EPA has identified these additional areas for Clean Air Scientific Advisory Committee (CASAC) review and comment.

<u>Ecological Effects of Ozone (Appendix 8)</u>

> *Please comment on the identification, evaluation, and characterization of the available scientific evidence from studies of ecological effects of ozone, and the application of information from these studies, as presented in Appendix 8 to inform causality determinations for these welfare outcomes.*

First of all, determinations are made about causation by evaluating evidence across scientific disciplines and are based on judgements of consistency, coherence and biological plausibility of observed effects, as well as related uncertainties. It was noted that the ISA used a formal causal framework to classify the "weight of the evidence" using a five level hierarchy that characterized the evidence that forms the basis of causality determinations for welfare effect categories of a "causal relationship" or a "likely to be causal relationship" or describe instances where causality determination has changed (i.e., "likely to be causal" changed to "suggestive of, but not sufficient to infer a causal relationship"). Other relationships between ozone and welfare effects include "suggestive of, but not sufficient to infer" and "inadequate".

There are 12 causality determinations for ecological effects of ozone that are generally organized from the individual-organism scale to the ecosystem scale presented in Figure ES-5 in the ISA. To summarize the findings of the 2013 Ozone ISA, five are causal relationships (i.e., visible foliar injury, reduced vegetation growth, reduced crop yield, reduced productivity, and altered

below ground biogeochemical cycles), and two are likely to be causal relationships (i.e., reduced carbon sequestration and altered ecosystem water cycling). One of the endpoints, alteration of terrestrial community composition, has now been concluded to be a "causal relationship" wherein the 2013 Ozone ISA this endpoint was classified as "likely to be causal". Three new endpoint categories (i.e., increased tree mortality, alteration of herbivore growth and reproduction, alteration of plant-insect signaling) not evaluated in the 2013 Ozone ISA are all determined to have a "likely to be causal relationship" with ozone. Plant reproduction, previously considered as part of the evidence for growth effects is now a stand-alone causal relationship as illustrated in Figure ES-5.

Visible foliar injury from ozone exposure has been well characterized and documented over decades involving many trees, shrubs, herbaceous and crop species in using both long-term field studies and laboratory approaches. Even more recent experimental evidence continues to show consistent association between visible injury and ozone exposure supporting a "causal relationship" between ozone and visible foliar injury. Consistent with the 2013 Ozone ISA, there is a "causal relationship" between ozone and reduced plant growth and a "causal relationship" between ozone and reduced crop yield and quality. In the 2013 Ozone ISA, EPA considered reproduction in the same category with plant growth. Increased information of plant reproduction (such as flower number, fruit number, fruit weight, seed number, rate of seed germination) and evidence for direct negative effects on reproductive tissues, as well as for indirect negative effects (resulting from decreased photosynthesis and other whole plant physiological changes) warrants a special causality determination of a "causal relationship" between ozone exposure and reduced plant reproduction. Since the 2013 Ozone ISA, large-scale statistical analysis of many factors concluded that county-level ozone concentrations averaged over the study period significantly increased tree mortality and many plant functional types. This evidence, combined with observations of long-term declines of conifer forests in several high ozone regions and new experimental evidence that sensitive genotypes of, particularly, aspen trees have increased mortality with ozone exposure, support a "likely to be causal relationship" between ozone exposure and tree mortality.

In addition to the direct effects of ozone on plants, ozone can alter ecological interactions between plants and other species, including herbivores that may consume ozone-exposed vegetation. Some recent evidence of insect herbivores in previous ozone assessments and new studies covering a range of species provide collective evidence that supports a "likely to be causal relationship" between ozone exposure and altered herbivore growth and reproduction. Many plant-insect interactions are mediated between volatile plant signaling compounds, which plants use to signal other members within an ecological community. New evidence from multiple studies show altered/degraded emissions of chemical signals from plants and reduced detection of plant signaling compounds by insects. Therefore, the collective evidence supports "a likely to be causal relationship" between ozone exposure and alteration of plant-insect signaling.

At the ecosystem scale, ozone caused suppression of plants' photosynthesis which can lead to reduced ecosystem carbon content. Consistent with the conclusions of the 2013 Ozone ISA, there is a "causal relationship" between ozone exposure and reduced productivity and a "likely to be causal" relationship between ozone and reduced carbon sequestration. Recent evidence continues to support a "causal relationship" between ozone exposure and the alteration of below ground biogeochemical cycles. We know ozone can affect water use in plants through several mechanisms and ultimately affect plant evapotranspiration, which may in turn lead to possible effects on hydrogeological cycling. Evidence continues to support the conclusion of the 2013 Ozone ISA that there is a "likely to be causal relationship" between ozone and alteration of ecosystem water cycling. Alteration of community composition of some ecosystems, including conifer forests, broadleaf forests and grasslands, and altered fungal and bacterial communities in the soil reported in the 2013 Ozone ISA is augmented by additional evidence for effects in forests and grassland communities indicating a change in the causality determination to a "causal relationship" between ozone exposure and altered terrestrial community composition of some ecosystems.

The summary of causality determinations for ecological effect are summarized as follows:

1. Conclusions from the 2013 Ozone ISA that support the seven conclusions of causality in the current 2019 Ozone ISA include 1) visible foliar injury, 2) reduced vegetation growth, 3) reduced plant reproduction, 4) reduced yield and quality of agricultural crops, 5) reduced productivity in terrestrial ecosystems, 6) alteration of below ground biogeochemical cycles, and 7) alteration of terrestrial community composition.

The "weight of the evidence" appears to strongly support the previous conclusions from the 2013 Ozone ISA subsequently identified in the conclusions in the current ISA. The summary of five causality determinations for ecological effects in the 2019 Ozone ISA, which build on the conclusions from the 2013 Ozone ISA, include the following:

1. Reduced plant reproduction from no "separate causality" to a "causal relationship" with ozone exposure,
2. Increased tree mortality "causality not assessed" and changed to "likely to be a causal relationship",
3. Alternation of herbivore growth and reproduction changed from "causality not assessed" to "likely to be causal relationship",
4. Alteration of plant-insect signaling "causality not assessed" changed to "likely to be a causal relationship",
5. Alteration of terrestrial community composition changed from "likely to be a causal relationship" to "causal relationship".

For these five causality determinations for ecological effects that have changed in terms of conclusions in the current ISA from the conclusions from the 2013 Ozone ISA will be more fully evaluated in terms of preliminary comments from the initial review of these data.

Appendix 8 "Ecological Effects" in the 2019 Ozone ISA evaluates the relevant scientific information on ecological effects as part of the review of the air quality criteria for ozone and other photochemical oxidants and to help form the scientific foundation for the review of the secondary NAAQS for ozone. This Appendix serves as an update to Chapter 9 of the 2013 Ozone ISA (U.S. EPA, 2013). The majority of the evidence for ecological effects has been for vegetation. Effects at the individual plant level can result in broad ecosystem-level changes, such as productivity, carbon storage, water cycling, nutrient cycling, and community composition. The current ISA has adopted the use of the Population, Exposure, Comparison, Outcome, and Study design (PECOS) tool to further define the scope of the current review by conveying the criteria for inclusion or exclusion of studies. The units of study as defined in the PECOS for ecological effects of ozone are the individual organism, species, population, community, or ecosystem. It should be noted that all studies included in the 2019 Ozone ISA were conducted at concentrations occurring in the environment or experimental ozone concentrations within an order of magnitude of recent concentrations observed in the U.S. For ecological endpoints for which the 2013 Ozone ISA concluded that the evidence was sufficient to infer a causal relationship (i.e., foliar injury, vegetation growth, ecosystem productivity, yield and quality of agricultural crops, below ground biogeochemical cycling). These were fully evaluated in the 2019 Ozone ISA. In terms of new determination or change in causality from the 2013 Ozone ISA, the following causality determinations for ecological effects of ozone will be addressed in the current review. At the community level, biodiversity in terms of terrestrial community composition is now "causal", and species interactions including plant-insect signaling is a new determination and "likely to be causal". In addition, tree survival is changed to "likely causal" and growth of insect herbivores feeding on ozone-affected plants is "likely causal". The plant reproduction endpoint is now separate from plant growth and a new determination as "causal" and new determination of growth and reproduction is "likely to be causal" is assigned to insect herbivores affected by ozone. All causality determinations or changes in causality determination from the 2013 Ozone ISA will be thoroughly considered in the present series of comments. The current review only evaluates studies conducted in North America. In the PECOS for ecological effects, relevant study designs include laboratory, greenhouse, field, gradient, open top chamber (OTC), free air carbon dioxide enrichment (FACE), and modeling studies.

<u>Visible Foliar Injury in Biomonitoring</u>

In the 2013 Ozone ISA, the evidence was sufficient to conclude that there is a causal relationship between ambient ozone exposure and the occurrence of ozone-induced visible foliar injury on sensitive plant species across the U.S. (U.S. EPA, 2013). Visible foliar injury from exposure to ozone has been well characterized and documented on many tree, shrub, herbaceous, and crop

species through research beginning in 1958. Ozone-induced visible foliar injury is considered diagnostic because it has been experimentally induced and it is considered a bioindicator for ozone exposure in plants. As described in the PECOS tool, the scope for new evidence reviewed in the section limits studies to those conducted in North America at concentrations occurring in the environment or experimental ozone concentrations within an order of magnitude of recent concentrations. Experimental evidence continues to show a consistent association between visible injury and ozone exposure in plants. Since the 2013 Ozone ISA, several studies have further characterized modifying factors 1) additional field studies have shown dry periods tend to decrease the incidence and severity of ozone-induced visible foliar injury, 2) data used in additional species from greenhouse studies add to the evidence that sensitivity to ozone varies by the time of day in plants, 3)phenotypic variation of foliar sensitivity to ozone has been observed, 4) in OTC exposure (mean 12 hour ozone concentration of 37 ppb for 118 days) foliar injury to loblolly pine seedlings were not related to seedling inoculation with root-infecting fungi (Chieppa et al, 2015).

Since the 2013 Ozone ISA, several additional studies have been conducted on bioindicator species:

1. Cutleaf coneflower is an ozone bioindicator species native to Great Smokey Mountains National Park,
2. Tree of heaven, an established invasive species found widely across the U.S., has been identified as an effective ozone bioindicator species by the National Park Service and Forest Service (Smith et al, 2008; Kohut, 2007).

In greenhouse exposures, foliar injury occurred at 8 hour average ozone exposure levels of 60 – 120 ppb with greater injury corresponding to higher exposure (Seiler et al, 2014). As noted in the 2013 ISA, visible foliar injury usually occurs when sensitive plants are exposed to elevated ozone concentrations in a predisposing environment. A major modifying factor for ozone-induced visible foliar injury is the amount of soil moisture available to a plant during the year that the visible foliar injury is being assessed. This is because the lack of soil moisture generally decreases stomatal conductance of plants and, therefore, limits the amount of ozone entering the leaf that can cause injury. Visible foliar injury from ozone exposure has been well characterized for decades using both long-term field studies and laboratory approaches. Since the 2013 Ozone ISA, new research on bioindicator species and the further characterization of modifying factors have provided further support for the effects. New information is consistent with the conclusions of the 2013 Ozone ISA that the body of evidence is sufficient to infer a causal relationship between ozone exposure and visible foliar injury. With the decades of research, both in field observation as well as experimental studies related to the foliar injury endpoint, the body of evidence remains very strong to infer a causal relationship between ozone exposure and visible foliar injury.

Plant Growth

In the 2013 Ozone ISA, the evidence was sufficient to conclude that there is a causal relationship between ambient ozone exposure and reduced growth of native woody and herbaceous vegetation (U.S. EPA, 2013). In the 2013 Ozone ISA, it was concluded there is strong and consistent evidence that exposure to ozone decreases photosynthesis and growth in numerous plant species. The evidence available at that time and now discussed in the 2019 Ozone ISA found that ambient ozone concentrations caused decreased growth (measured as biomass accumulation) in annual, perennial, and woody plants inclusive of crops, annuals, grasses, shrubs, and trees. A meta-analysis by Wittig, et al (2009) found that the average ozone exposures of 40 ppb significantly decreased annual total biomass by 7% across 263 studies. Biomass declines were linked to reductions in photosynthesis (U.S. EPA, 2013), which are consistent with cumulative plant uptake of ozone into the leaf (Wittig et al, 2007). Further, there is evidence ozone may change plant growth patterns by significantly reducing carbon allocated to roots in some species. Since the 2013 Ozone ISA, there is more evidence from experimental studies that support detrimental effects of ozone on plant growth:

1. Results from aspen-only stand at the Aspen FACE experiment in Wisconsin showed a decrease of 12 – 19% in the relative growth rate of 3 of 5 genotypes of aspen studied,
2. When site level results from Aspen FACE experiment were scaled up using the forest landscape model (LANDIS II), ozone was found to significantly reduce landscape biomass,
3. In meta-analysis of 9 studies examining intra-specific variation in juvenile tree growth under elevated ozone, found that elevated ozone generally reduced photosynthetic rate as well as height growth and stem volume,
4. A study using the invasive Chinese tallow tree suggests ozone response may be genotype-specific,
5. Using model simulation coupled with established U.S. EPA ozone exposure response functions in seedlings, estimated relative biomass loss at 2.5% for Ponderosa pine and 2.9% for aspen, and
6. In another estimation of biomass loss of adult trees across the U.S. for modeled ozone values, eastern cottonwood and black cherry showed high sensitivity.

In addition to these studies, there is a recent global scale synthesis of published ozone exposures studies that document reductions in biomass due to ozone exposure in over 100 plant species (Bergmann et al, 2017). In the 2019 Ozone ISA, there is strong scientific evidence sufficient to conclude that there is a causal relationship between ambient ozone exposure and reduced growth of native woody and herbaceous vegetation.

Reduced Plant Reproduction

In the 2013 Ozone ISA, reduced plant reproduction was not separated for causality determination and was included with plant growth. However, in the 2019 Ozone ISA, reduced plant reproduction is scientifically defended for a causal relationship between plant reproduction metrics and exposure to ozone. In fact, the recent literature shows that across most plant reproduction metrics (such as flower number, fruit number, fruit weight, seed number, and rate of seed germination) with elevated exposure concentrations that ozone has significant negative effects on plant reproduction. In a first of its kind study, Leisner and Ainsworth (2012) conducted a quantitative meta-analysis to assess the general magnitude and direction of the effects of ozone exposure on plant reproduction. In experiments that used ambient air as the control, average fruit weight decreased 51% (at an average exposure of 98 ppb), which was the largest effect observed in this part of the meta-analysis, and seed number decreased approximately 10% (at an average exposure of 68 ppb). In studies with ozone-sensitive species of clover, Sanz et al (2016) showed that reproduction was reduced significantly with increasing ozone exposure. Gillespie et al (2015) isolated the effects of ozone on particular reproductive tissues of tomato. Pollen grains exposed to ozone have significantly reduced germination and pollen tube growth in vitro. Reductions in pollen viability is, for example, and extremely important plant reproduction metric.

Timing of ozone exposure relative to reproductive development stages can affect reproductive outcomes in some cases. Flowers exposed to ozone early in their development tended to produce shorter fruits than flowers exposed later in their development. There appears to be adequate information, particularly from the quantitative meta-analysis reported by Leisner and Ainsworth (2012) supporting a causal relationship between ozone exposure and reduced plant reproduction. The strength of the scientific support for supporting a “causal relationship” is not as strong as with visible foliar injury and reduced vegetation growth as previously reviewed in the current comments. However, with the separate category of reduced plant reproduction, it can be concurred that causality does exist between ambient ozone exposure and this plant metric. It has been shown that diverse metrics of plant reproduction decline under ozone concentrations occurring in either the environment or under experimental conditions within an order of magnitude of recent concentrations. Metrics of plant reproduction, fruit number and fruit weight, show reductions under increased ozone when combined across species for ozone concentrations that span 40 to >100 ppb. Finally, experimental ozone exposure at multiple experimental settings (such as in vitro, whole plants in the laboratory, whole plants and/or reproductive structures in the greenhouse, and whole plant communities in field settings) convincingly show ozone independently reduces plant reproduction. I concur that previous evidence and new evidence reviewed here is sufficient to infer a “causal relationship” between ozone exposure and reduced plant reproduction.

Plant Mortality

In the 2013 Ozone ISA, causality was not assessed for increased tree mortality involving ozone exposure. The conclusions in the 2019 Ozone ISA is that there is "likely to be a causal relationship" between ozone exposure and plant mortality. Several new studies examine the impacts of ozone exposure on plant mortality that included the fraction of individuals in a population that die over a given timeframe. These experiments were focused on tree species demonstrating ozone exposure can affect tree mortality. For instance, in the Aspen FACE experiment, the survival of sensitive aspen genotypes 271 and 259 declined significantly between 1997 and 2008 under elevated ozone exposures (Moran and Kubiske, 2013). In addition, Dietz and Moorcroft, 2011) conducted a large-scale analysis of factors contributing to annual mortality of trees and functional types in the forests of the eastern and central U.S. In their analysis, ozone was ranked 9$^{th}$ on a list of 13 factors that forests were sensitive to an ozone's effects with similarly magnitude to that of precipitation. Mortality in 8 out of 10 plant functional types were significantly correlated with ozone 8 hour max exposures. Therefore, studies of tree mortality indicate that ozone affects this endpoint. Studies linking ozone and tree mortality are consistent with known and well-established individual plant level mechanisms that explain ozone phytotoxicity, including variation and sensitivity and tolerance based on age class, genotype, and species. Experimentally elevated ozone exposures has been shown to increase mortality in sensitive Aspen genotypes. Considering the previous evidence and new evidence reviewed in the 2019 Ozone ISA, it is sufficient to infer a "likely to be causal relationship" between ozone exposure and tree mortality.

Reduced Crop Yield and Quality

In the 2013 Ozone ISA, the evidence was sufficient to conclude there is a "causal relationship" between ozone exposure and reduced yield and quality of agricultural crops (U.S. EPA, 2013). The detrimental effect of ozone on crop production has been recognized since the 1960s and there is a large body of research that has subsequently characterized decreases in yield and quality of agricultural crops. As described in the PECOS tool, the scope of new evidence reviewed in this section are limited to studies conducted in North America at ozone concentrations occurring in the environment or experimental ozone concentrations within an order of magnitude of research concentrations.

For soybeans, additional studies in Illinois report decreased seed/crop yield (Leisner et al, 2017). A linear decrease in soybean yield was observed across two growing seasons at the rate of 37-39 kg/ha per ppb cumulative ozone exposure over 40 ppb. For wheat, meta-analysis using data from the U.S. and other countries provide further supporting evidence that current levels of ambient ozone decrease growth, quality, and yield (Pleijel et al, 2018). New studies in non-soybean legumes include evaluation of biomass and seed yield in ozone-exposed snap bean under high- and low-vapor pressure deficit conditions (Fiscus et al, 2012). U.S. modeling studies in the 2013

Ozone ISA found that ozone generally reduced crop yield and that different crops showed different sensitivity to ozone (Avnery et al, 2011). Newly available regional and national scale analyses of ozone effects on major crops in the U.S., including soybean, wheat, and maize have further enabled characterization and quantification of yield losses (McGrath et al, 2015).

The relationship between ozone exposure and reduced crop yield is well established in the scientific literature and continues to be an active area of research with many new scientific papers being published since the 2013 Ozone ISA. Recent advances in characterizing ozone effects on U.S. crop yield include further geographic and temporal refinement of ozone sensitivity in national scale estimates of maize and soybean losses from ozone based on actual yield data. The new scientific information published is consistent with the conclusions of the 2013 Ozone ISA that the body of evidence is sufficient to infer a "causal relationship" between ozone exposure and reduced yield and quality of agricultural crops.

<u>Herbivores: Growth, Reproduction, and Survival</u>

In the 2013 Ozone ISA, there was no causality determination between ozone exposure and effects on herbivores. We know that ozone exposure can lead to changes in plant physiology, such as by modifying the chemistry and nutrient content of leaves (U.S. EPA, 2013). These changes can have significant effects on herbivore physiology and behavior. There was no consensus in the 2013 Ozone ISA on how insects and other wildlife respond to elevated ozone. Since that review, additional research has been published for more herbivorous insects as well as a few mammalian herbivores at various levels of ozone exposure. As described in the PECOS tool, the scope of this review includes studies in which alterations in invertebrates and vertebrate responses were measured in individual species or at the population and community levels as related to concentrations of ozone occurring in the environment or experimental ozone concentrations within an order of magnitude of recent concentrations. In the 2013 Ozone ISA, a meta-analysis that included 16 studies published on insect herbivore species between 1996 and 2005 found that elevated ozone decreased development time and increased pupil mass in insect herbivores with more pronounced effects occurring with longer durations of ozone exposure (Valkama et al, 2007). Since the 2013 Ozone ISA, there is new evidence for endpoints related to growth, reproduction, and survival in insect herbivores encompassing the orders Coleoptera, Hemiptera, and Lepidoptera. With the available science reported in the 2019 Ozone ISA regarding the effects of ozone on growth, reproduction, and survival of, particularly, insect herbivores substantial new information has been made available in order to assess a causality relationship. In addition, population and community level responses reveal that changes in host – plant quality resulting from elevated ozone can alter the population density and structure of associated insect herbivore communities ultimately affecting ecosystem processes (Cornelissen, 2011). Recent studies reviewed in the 2019 Ozone ISA include multiple experimental studies conducted by many research groups that expand the evidence base for the effects of elevated ozone on growth and reproduction in herbivores. It is recognized that while effects were

observed there remains a more limited number of studies on the effects of ozone on survival and population/community level responses. Recognizing that since the 2013 Ozone ISA and with increased research efforts on herbivore response to plants impacted by ozone, a new causality determination appears justified that the body of evidence is sufficient to infer a "likely to be causal" relationship between ozone exposure and alteration of herbivore growth and reproduction.

Alteration of Plant-Insect Signaling

In the 2013 Ozone ISA, there was "no causality" determination between ozone exposure and alteration of plant-insect signaling. Plants signal to other ecological community members through the emission of volatile plant signaling compounds (Blande et al, 2014). Each signal emitted by plants has an atmospheric lifetime and a unique signature comprised of different ratios of individual hydrocarbons that are susceptible to atmospheric oxidants, like ozone (Yuan et al, 2009). Insects and other fauna discriminate between chemical signals of different plants. As described in the PECOS tool, the scope in the 2019 Ozone ISA for considering plant-insect signaling include studies that assess altered plant-insect signaling in response to concentrations of ozone occurring in the environment or experimental ozone concentrations within an order of magnitude of recent concentrations. Under conditions of elevated ozone the degradation of plant signaling compounds resulted in bumble bees orienting significantly less towards floral scent queues and exhibiting preference for artificial flowers closer to the ozone source (Farre-Armengol et al, 2015). As reported previously, herbivorous insects use plant signaling compounds to locate suitable host plants and ozone can alter these interactions (Blande et al, 2010). In chamber studies, elevated ozone reduced the ability of insect herbivores to find their plant host (Li et al, 2016). Striped cucumber beetles could not distinguish between clean air and air containing floral volatiles when the ozone concentration exceeded 80 ppb (Fuentes et al, 2013). In addition, plant defense responses include emission of plant-signaling compounds to attract predators and parasitoids that target the herbivores feeding on the plant. In studies reviewed in the 2013 Ozone ISA and new studies on parasitoid-host attraction show either reduced, enhanced, or unaffected behavior by elevated ozone (Cui et al, 2016). Altered plants signaling to natural enemies of herbivores disrupts predator-prey trophic interactions. The interaction of ozone (>50 ppb) with plant signaling compounds disrupts the production, emission, dispersion, and lifespan of these compounds. Considering the available evidence reported in the 2013 Ozone ISA and more recent research efforts while as well recognizing uncertainties around how chemical signaling responses observed in the laboratory translate to natural environments, the 2019 Ozone ISA makes a new causality determination that the body of evidence is sufficient to infer a "likely causal relationship" between ozone exposure and alteration of plant-insect signaling.

Reduced Productivity in Terrestrial Ecosystems

In the 2013 Ozone ISA, the evidence was sufficient to conclude there is a causal relationship between ozone exposure and reduced plant productivity. The terrestrial carbon cycle integrates processes at various scales, ranging to organelles to individuals to biomes (Chapin et al, 2002). Gross primary productivity, which is the influx of $CO_2$ from the atmosphere via photosynthesis at the ecosystem scale is fundamental to global carbon cycling. Since the 2013 Ozone ISA, two new studies have reported on the effects of ozone on gross primary productivity. Fares et al (2013) conducted statistical analysis of data to quantify the effect of ozone on carbon assimilation. In California, ozone decreased carbon assimilation by 12% in pine forests in the Sierra Nevada and by 19% in an orange grove in the Central Valley. Yue and Unger (2014) adopted the same ozone-damaged thresholds in their analysis that were used in previous models to assess ozone damage. What was learned was decreases in gross primary productivity as a result of ozone range from 1-14% and were greatest at sites showing both high stomatal conductance and high growing season ozone concentrations. Carbon assimilated into plant tissue via photosynthesis is either respired or contributes to net primary productivity, which is often measured as the rate of plant biomass accumulation. While much of the research published since 2013 Ozone ISA is confirmatory, other work has provided new mechanistic insight into the effects of ozone on net primary productivity. Evidence of the effect of ozone exposure in ecosystem productivity comes from many different experiments with different study designs in a variety of ecosystems and models. New information is consistent with conclusions of the 2013 Ozone ISA that the body of evidence is sufficient and increasing to infer a "causal relationship" between ozone exposure and reduced ecosystem productivity.

Reduced Carbon Sequestration in Terrestrial Ecosystems

Terrestrial carbon sequestration is the sum of carbon contained within biomass and soil within a defined ecosystem typically quantified on a multi-year scale (Koerner 2006). As in the 2013 Ozone ISA, most assessments of the effects of ozone on terrestrial carbon sequestration are from model simulations. However, an assessment of the effect of ozone on ecosystem carbon content at the Aspen FACE experiment was published in 2014. At the conclusion of the Aspen FACE experiment after 11 years of fumigation, Talhelm et al (2014) observed that elevated ozone decreased ecosystem carbon content (in plant biomass, litter, and soil carbon to 1 m in depth) by 9%. Total tree biomass carbon was 15% lower under elevated ozone with decreased woody biomass counting for nearly all of the effect of tree biomass. The results from the Aspen FACE experiment and the model simulations provide further evidence that ozone can decrease ecosystem carbon sequestration. Although the decrease in net primary productivity were temporary in the Aspen FACE experiment, the 10% decrease in cumulative net primary productivity at Aspen FACE was associated with a 9% decrease in ecosystem carbon storage (Talhelm et al, 2014). The relationship between ozone exposure and terrestrial carbon sequestration is difficult to measure at the landscape scale. Most of the evidence regarding this

relationship is from model simulations, although this endpoint was examined in a long-term manipulative chamber-less ecosystem experiment known as Aspen FACE, already described. Even with limitations, the result from the Aspen FACE experiment and supported by model simulation provide further evidence that is consistent with the conclusions of the 2013 Ozone ISA that the body of evidence is sufficient to conclude there is a "likely to be causal relationship" between ozone exposure and reduced carbon sequestration in ecosystems.

Soil Biogeochemistry

The 2013 Ozone ISA concluded there is a "causal relationship" between ozone exposure and the alteration of below ground biogeochemical cycles (U.S. EPA, 2013). This causality determination was based on the body of evidence known at that time. The 2013 Ozone ISA (U.S. EPA, 2013) presented evidence that ozone alters multiple below ground endpoints, including root growth, soil food web structure, soil decomposer activities, soil respiration, soil carbon turnover, soil water cycling, and soil nutrient cycling. The new evidence since the 2013 Ozone ISA included in the 2019 Ozone ISA confirms ozone affects soil decomposition, soil carbon, and soil nitrogen. Soil carbon is often a mix of inorganic and organic forms of carbon, the latter may be from living and/or dead plant animal, fungal, and bacterial organisms. The effects of ozone on several aspects of soil carbon have been investigated. Ozone can alter the cycling of nitrogen in the soil via its direct effect on plants. Nitrogen is an important element to plant life as it is often the limiting nutrient from most temperate ecosystems. The 2013 Ozone ISA (U.S. EPA, 2013) documented mixed results of ozone effects on soil nitrogen pools and processes with results indicating no effect in meadow nitrogen biomass or potential nitrification and denitrification (Kanerva et al, 2006). While ozone was shown to increase nitrogen released from litter in a forest (Stoelken et al, 2010), ozone decreased gross nitrogen mineralization (Holmes et al, 2006) at Aspen FACE and nitrogen release from soil litter. The 2013 Ozone ISA presented evidence that ozone was found to alter multiple below ground endpoints, including root growth, soil food web structure, soil decomposer activities, soil respiration, soil carbon turnover, soil water cycling, and soil nutrient cycling. New evidence since the 2013 Ozone ISA (U.S. EPA, 2013) included in this assessment confirms ozone effects on soil decomposition, soil carbon, and soil nitrogen. Overall, the evidence does not change the conclusions from the 2013 Ozone ISA (U.S. EPA, 2013) and, therefore, suggests that ozone can alter soil biogeochemical cycling of carbon and nitrogen, although the direction and magnitude of these changes often depends on the species, site, and time of exposure. Currently, it is recognized that it does not appear to be a consistent exposure-response relationship. The body of evidence is sufficient to conclude that there is a "causal relationship" between ozone exposure and the alteration of below ground biogeochemical cycles.

Alteration of Terrestrial Community Composition

In the 2013 Ozone ISA, the evidence was sufficient to conclude there is a "likely to be causal relationship" between ozone exposure and alteration of terrestrial community composition of

some ecosystems (U.S. EPA, 2013). Ozone altered above ground plant communities, such as conifer forests, broadleaf forests, and grasslands and altered fungal and bacterial communities in the soil in both natural and agricultural systems (U.S. EPA, 2013). Ozone effects on individual plants can alter the larger plant community as well as the below ground community of microbes and invertebrates, which depend on plants as carbon sources. In the 2013 Ozone ISA, evidence of ozone effects on forest composition was drawn from the observational studies of conifer decline correlated with ozone exposure (Allen et al, 2007). New evidence suggests that ozone alters tree competitive interactions for nutrients, such as consistent with previous research on altered tree community composition at Aspen FACE showed that elevated ozone altered the relative competition for nutrients among aspen genotypes (Zak et al, 2012). Since the 2013 Ozone ISA, new studies extend the scope of evidence regarding forest community composition to include synthesis and models. In the 2013 Ozone ISA, there was evidence of ozone effects on grassland community composition in controlled experimental exposure studies, in models, and in reviews. Key new studies include experimental ozone exposures that allow evaluation of ozone effects on grassland community composition and analyses that explicitly include environmental or annual heterogeneity.

Even with microbes, the 2013 Ozone ISA documented effects of ozone on soil microbial communities with changes in proportions of bacteria or fungi as a result of experimental ozone exposures in grassland mesocosms, peatland mesocosms, and forest mesocosms. In addition, changes in soil microbial communities in agricultural systems was reported (Chen et al, 2010). Even with bacteria, the 2013 Ozone ISA found decreases in bacterial abundance in response to elevated ozone in meadows and forests mesocosms. There have been many new studies reported to assess the effect of elevated ozone on soil bacteria. The 2013 Ozone ISA found effects of ozone exposure on soil fungi (U.S. EPA, 2013). Studies found that ozone exposure decreased fungal biomass in meadow mesocosms, marginally increased fungal abundance in peatland mesocosms and altered fungal community composition in forest soils. Many new studies have evaluated the effects of ozone on fungi since the 2013 Ozone ISA. The 2013 Ozone ISA found evidence sufficient to conclude that there is a “likely to be causal relationship” between ozone exposure and the alteration of community composition of some ecosystems. Evidence of this relationship was presented for forest communities of trees, grassland communities of grasses, herbs, and legumes and soil microbial communities of bacteria and fungi. Recently published papers extend the evidence of each of these topics in the 2013 Ozone ISA.

In forests, previous evidence included correlation on studies across ambient gradients of ozone exposure that found effects of ozone on conifer trees, studies with controlled experimental exposure of trees that found effects of ozone on deciduous trees. Key new studies show that observational and experimental observations of ozone effects on tree species extend to affect regional forest composition in the Eastern U.S. (Wang et al, 2016). In grasslands, previous evidence included multiple studies from multiple research groups to show that elevated ozone shifts the balance among grasses, forests, and legumes. There are new studies that show ozone

affected the ratio of grass to legume biomass (Gilliland e al, 2016). In soil microbial communities, previous evidence includes studies that found effects on the ratio of bacteria to fungi in soil communities as well as effects on community composition of mycorrhizal fungi. New studies confirm that elevated ozone alters soil microbial taxa, although as with previous evidence, the strength and directional effects are not consistent across all ecosystems. The 2013 Ozone ISA presented multiple lines of evidence that elevated ozone alters terrestrial community composition, and recent evidence strengthens our understanding of the effects of ozone on plant communities while confirming that the effects of ozone on soil microbial communities are diverse. The body of evidence is sufficient to conclude that there is a "causal relationship" between ozone exposure and the alteration of community composition of some ecosystems.

Alteration of Ecosystem Water Cycling

In the 2013 Ozone ISA, the evidence was sufficient to conclude there is a "likely to be causal relationship" between ozone exposure and the alteration of ecosystem water cycling (U.S. EPA, 2013). Plants are responsible for part of the ecosystem water cycling through root uptake of soil moisture and groundwater as well as transpiration through leaf stomata to the atmosphere. Changes to this part of the water cycle may in turn affect the amount of water moving through the soil, running off over land or through groundwater and flowing through streams. Ozone can affect water use in plants and ecosystems through several mechanisms, including damage to stomatal functioning and loss of leaf area, which may affect plant and stand transpiration. During the review of the 2013 Ozone ISA, there was debate on the assumption that ozone exposure consistently reduced stomatal conductance in plants. Several studies have found increased conductance, suggesting stomatal dysfunction in response to ozone exposure. However, other studies found ozone caused a loss of stomatal control, incomplete stomatal closure at night, and a decoupling of photosynthesis in stomatal conductance. There is mounting biologically relevant and statistically significant data from multiple studies showing the mechanisms of ozone effects on plant-water use in ecosystem water cycling (reduced leaf area, reduced leaf longevity, changes in root and branch biomass and architecture, changes in vessel anatomy, stomatal dysfunction, reduced sap flow). The most compelling evidence showing effects at the ecosystem level is from studies in Eastern U.S. forests and from the Aspen FACE. All of this new information supports the 2013 Ozone ISA and supports the conclusion in the 2019 Ozone ISA that the body of evidence is sufficient to conclude there is a "likely to be causal relationship" between ozone exposure and the alteration of ecosystem water cycling.

General Comments

1. I compliment the United States Environmental Protection Agency for the thoroughness and completeness of Appendix 8 entitled, "Ecological Effects" as part of the 2019 Ozone Integrated Science Assessment.

2. I agree with the "causality" determinations for the components for ecological effects considered in the 2019 Ozone Integrated Science Assessment including 1) visible foliar injury, 2) reduced vegetation growth, 3) reduced plant reproduction, 4) reduced yield and quality of agricultural crops, 5) reduced productivity in terrestrial ecosystems, 6) alteration of below ground biogeochemical cycles, and 7) alteration of terrestrial community composition. I agree with the "likely to be causal" determinations including 1) increased tree mortality, 2) alteration of herbivore growth and reduction, 3) alteration of plant-insect signaling, and 4) reduced carbon sequestration in terrestrial ecosystems, and 5) alteration of ecosystem water cycling
3. It is my impression that a thorough review and reporting of the scientific literature that has been generated since the 2013 Ozone ISA has been incorporated into the 2019 Ozone ISA.
4. In terms of the summary of causality determinations for ecological effects, I support the determinations made by the U.S. EPA as a function of the science availability and its interpretation.
5. Although historically the predominant ecological effects assessed with ozone exposure has been with vegetation, the current Appendix 8 "Ecological Effects", has at least some mention of terrestrial vertebrates, including rabbits and goats, and how they may respond to altered vegetation as a function of ozone exposure. I think that this area should be expanded in terms of alteration of individual plants and plant communities can disrupt terrestrial vertebrates, and not just invertebrates. Therefore, I recommend consideration of an expanded research plan to look at the implications of altered vegetation communities from ozone exposure and response to terrestrial vertebrate herbivores.
6. Although there is in depth consideration in other sections of the 2019 Ozone Integrated Science Assessment involving human health implications from ozone exposure, which are real and well-defined cause and effect relationships that have been scientifically studied a considerable length of time, nothing is mentioned with wildlife. In Appendix 8 "Ecological Effects", there is no mention whatsoever of wildlife toxicology implications for ozone exposure, although human health implications have been considerably considered in other parts of the 2019 Ozone Integrated Science Assessment. I recommend to at least consider and develop a research plan for a bird model that could be assessed in terms of the wildlife toxicology of ozone exposure in warm-blooded vertebrates. This would be essentially a "canary in the coal mine" concept for detecting toxic gasses by miners through a bird model. I think this same concept could be implemented utilizing an avian model for the study of ozone exposure in terrestrial warm-blooded non-human vertebrates (Kendall at el, 2010).

<u>Climate Change (Appendix 9)</u>

*Please comment on the identification, evaluation and characterization of the available scientific evidence from studies of ozone effects on climate, and the application of*

*information from these studies as presented in Appendix 9 to inform causality determinations for these welfare outcomes.*

For effects on climate, changes in the abundance of tropospheric ozone disturbs the radiative balance of the atmosphere by interacting with incoming solar radiation and outgoing longwave radiation. This effect is quantified by radiative forcing, which is the perturbation in net radiation flux at the tropopause caused by a change in radiatively active forcing agent after stratospheric temperatures have readjusted to radiative equilibrium. Through this effect on the earth's radiation balance, tropospheric ozone plays a significant role in the climate system and increases in tropospheric ozone abundance contribute to climate change as addressed in the 2013 Ozone ISA. Recent evidence continues to support a causal relationship between tropospheric ozone and radiative forcing and a "likely to be causal relationship" via radiative forcing between tropospheric ozone and temperature, precipitation and related climate variables referred to as climate change in the 2013 Ozone ISA. New evidence comes from the Intergovernmental Panel on Climate Change (IPCC) Fifth Assessment Report (AR5) (Myhre et al, 2013) and supporting references. As thoroughly discussed in the 2019 Ozone ISA, none of the new studies indicate a change to either causality determination included in the 2013 Ozone ISA. In terms of effects of tropospheric ozone and climate change, radiative forcing remains a "causal" relationship and temperature, precipitation, and related variables maintain a "likely to be causal" relationship. Consistent with previous estimates in the 2013 Ozone ISA, the 2019 Ozone ISA is consistent with previous estimates, the effect of tropospheric ozone on global surface temperature through its impact on radiative forcing continues to be estimated at roughly 0.1 to 0.3°C since industrial times. While the warming effect of tropospheric ozone in the climate system is established, precisely quantifying changes in surface temperature due to tropospheric ozone changes along with related climate effects requires complex climate simulations. There are current limitations in climate modeling tools that need to be recognized and the need for more comprehensive observational data on these effects represent sources of uncertainty in quantifying the precise magnitude of climate response to ozone changes (Myhre et al, 2013). All of this evidence reinforces the "likely to be causal" relationship between tropospheric ozone and temperature, precipitation, and related climate variables which was referred to as "climate change" in the 2013 Ozone ISA.

<u>General Comments</u>

1. I compliment the United States Environmental Protection Agency for continuing to clearly characterize and communicate the effects of ozone as related to climate change building on the 2013 Ozone ISA to the current document, draft 2019 Ozone ISA.
2. Although evidence has increased supporting the relationship between tropospheric ozone and aspects of climate change, including a "causal relationship" with radiative forcing as well as a "likely to be causal relationship" with impacts on temperature, precipitation, and related climate variables, the causality determinations reached in the 2013 Ozone

ISA are even further supported in the 2019 Ozone ISA, and I strongly concur with that position.

3. Further research would be useful, particularly quantifying the relationship between regional ozone RF including from ozone aerosols and other short-lived climate forcers on the hydrologic cycle, precipitation, and atmospheric circulation patterns; improving understanding of and ability to model critical ozone-climate processes; and continuing exploration of links between precursor pollutant control strategies, climate, and ozone concentrations. These research strategies would be extremely useful as we continue to better understand the role of ozone in the climate system scientific arena.

# Dr. Corey Masuca

**Appendix 1 – Atmospheric Source, Chemistry, Meteorology, Trends, and Background Ozone**

1.3.1.2 Global and International Sources of Anthropogenic Ozone Precursors

While global and international sources of anthropogenic ozone precursors is adequately discussion, there appears to be a noticeable absence discussion(s) of localized, interstate and/or intercity anthropogenic ozone precursors.

1.5.1 Meteorological Effects on Ozone Concentrations at the Ground Level

There appears to be a noticeable absence of discussion(s) about topographical effects on/of ozone formation and ozone concentration transport (i.e., "bowl" inversion effects in valleys and areas surrounded by mountains).

**Appendix 2 – Exposure to Ambient Ozone**

2.4.2 Infiltration

While this section focuses on infiltration of outdoor ozone concentrations through/into indoors for home and buildings, have there been studies of ozone infiltration into vehicles, especially since most humans spend a significant amount of time in their vehicles during daily commutes?