Preliminary Comments from Members of the CASAC NOx Panel on EPA's *Integrated Review Plan for the Primary National Ambient Air Quality Standards for Oxides of Nitrogen Volume 2: Planning for the Review and the Integrated Science Assessment (Draft - March 2024)*

Received as of 4-10-24

Mr. Ed Avol

The IRP Plans provided describe a thoughtful and detailed process by which EPA will undertake and develop a relevant and complete Integrated Scientific Assessment (ISA) for Nitrogen Oxides. I list a few brief comments and questions below…

Volume 1:

(explanation of Administrator's justification for no-change-to-standard in 2018 review difficult to follow)

Volume 2:

P2-3, end of 1st paragraph, " As in past reviews, …' – it is entirely appropriate to continue this approach with regard to previously-identified susceptible sub-groups, but now that there has been several years of near-roadway data collected and studies performed with populations in those locations, consideration of vulnerable populations (by location, exposure, socio-economic status) should also be emphasized and a focus of the review.

P2-4, mid-page (2nd dash heading, 4th bullet) - I welcome and strongly endorse the consideration of multi-pollutant exposure information as a more realistic reflection of true-life exposures. How does the Agency envision incorporating this exposure reality – pre-exposure to other pollutants or concurrent co-exposures - into their assessment of NOx exposures (from epi OR clinical studies research)?

P2-4, bottom third of page (3rd dash heading, 2nd bullet) and – with regard to co-occurring risk factors, these CASAC reviews take place every 5 years or so, and there are substantial environmental and ecological changes currently underway (climate change, rapid electrification of some societal sectors leading to changes in energy generation, proliferation of wildfire smoke leading to downwind exposure of vulnerable populations, …), so how can or does this help to inform Agency thinking planning, and actions regarding standards setting?

P2-4 to P2-5 questions – has new information altered our understanding of the health effects of ambient NOx exposures (acute or chronic) in organ systems other than the heart or lungs (such as the brain, CNS, metabolic, kidneys, or liver)?

P3-4, Section 3.2.1 Health Effects – the discussion explains that peer-reviewed published "…studies meeting all five aspects of the PECOS statement (discipline-specific **p**opulation, **e**xposure, **c**omparison, **o**utcome, **s**tudy design) will be considered for inclusion in the ISA." In the June 9, 2023 letter from the CASAC to the EPA Administrator (EPA-CASAC-23-002) regarding the 2022 reconsideration of the ozone standard, the CASAC unanimously recommended that the PECOS criteria be broadened to include studies conducted outside the US and North America (see p3, para3 of the letter). Will this broadened application of PECOS be incorporated for use during this cycle of the NO2 review? Footnote #16 at the bottom of P3-5 is silent on this point…

P3-4, last sentence – "…The ISA will also integrate previous information on the populations and life stages at increased risk with new evidence for existing *and any newly identified risk factors*. Will staff consider climate change into this in some manner?

P3-16, Footnote 24 – why is there a differentiation in exclusion criteria for North American or non-North American populations? Isn't it possible that non-North American population studies could be useful for evaluating potential policy options?

P3-20, last bullet, "…how have emissions and concentrations of NOx and of NO2 changed since the 2016 ISA?" -in the past several years, wildfire smoke and downwind impacts have become more frequent and a source of exposure and health concern; what is the Agency's thinking about this reality, and how might this be integrated into better public health protection?

P3-29, Table 3-9: is there evidence for potential NO2 exposure and increased risk of NO2-related health effects downwind of urban areas due to photochemistry, or downwind of forested areas due to wildfire smoke?

P3-30: does evidence exist to address increased risk for pregnant mothers or newborn infants with respect to residence or day-care coverage near major roads, or during episodes of wildfire smoke?

General Questions:

1) Are mining operations considered "ambient" exposures, or are those closed-space exposures? How are these sorts of occupational exposures captured and integrated into the review?

Dr. Howard Chang

Page 3-15: in Evaluation of Individual Study Quality, it is noted that there is a focus on "validated models" used to estimate exposures "for the study locations and populations." Epidemiologic studies are increasingly using fusion air quality data products that are developed to achieve complete spatial and temporal coverage. However, these data products are often not created with a specific health study in mind, and validation with measurements remains a challenge in regions with sparse monitoring.

Page 3-29: in describing the list of potential at-risk populations and lifestages, should pregnant people be considered as a factor to be evaluated?

Page A-21: in evaluating short-term exposure health studies, it is noted that studies utilizing a case-crossover design will be emphasized (presumably over the time-series design). The justification of this statement is somewhat unclear. The case-crossover design also has several analytic challenges (e.g., requiring more complete daily exposure, the inability to account for dispersion, and the standard within-month self-controls can be insufficient for respiratory outcomes).

Page A-22: in the discussion of study design, it is noted that recent studies have employed more diverse and flexible methods to handle measured and unmeasured confounders. It is appreciated that these studies will also be carefully evaluated for their assumptions.

Page A-27: for the list of important confounders for short-term health effects, would medication use that varies temporally be a possible mediator for air pollution effects? For long-term health effects, are historic sources referring other chronic environmental co-exposures?

Page 3-8: it is noted that the impact of the COVID-19 pandemic on emissions and ambient air quality will be discussed. Is this restricted to the short lockdown periods? It may be helpful to describe the associated policy relevant objectives of these discussions. Similar to the health studies (Section 3.3.3), some possible challenges (e.g., difficulty to disentangle meteorological effects in this short period) can also be included.

Minor point of clarification: mortality outcome is referred to as "total" mortality throughout the IRP. Does total include all-cause (non-accidental) and cause-specific deaths?

Dr. Deborah Cory-Slechta

My only comment relates to the following:

> "The health chapters of the ISA will evaluate the scientific literature related to a range of health outcomes associated with exposures to oxides of nitrogen including, but not limited to, respiratory effects, cardiovascular effects, reproductive and developmental effects, cancer, and mortality."

Albeit the text indicates "not limited to," I think it is important to include neurological effects in this category. The scientific literature relating air pollution to effects on brain has increased dramatically over the past 15 years. While not all studies include measures of oxides of nitrogen, it is nevertheless important to examine those that do given the potential to regulate this component of air pollution should it be found to be contributory to neurological effects.

Mr. Dirk Felton

Section 3.2.1

PECOS exposure criteria for epidemiologic studies are provided for annual averages in comparison to the 98th percentile of annual average NO_2 concentrations measured at ambient air monitors in the U.S. NO_2 is seasonal in some parts of the country so is there consideration for studies where there are higher wintertime concentrations?

Are there any other concentration or averaging period restrictions on studies looking at the annual or shorter-term exposures? At urban and near-road sites, higher concentrations typically occur for 4-5 hours on weekday mornings.

Section 3.2.2

After 2010, the EPA changed the quality assurance priority from accuracy around the level of the NAAQS to accuracy around the middle of the routine concentrations on a site-specific basis. Accuracy at low NO_2 levels is difficult due to drift and the influence of humidity on instrument stability. Is this an appropriate strategy for the State and Local monitoring agencies?

Section 3.4.1

Wildfire smoke has more NO than NO_2 depending on age so the consideration of the indicator should take this into account.

Dr. Christina H. Fuller

On page 2-4 of the document the consideration of new information "regarding the human life stages and populations that are particularly at increased risk for experiencing health effects" is described. I suggest the EPA consider distinguishing this information in three categories: external vulnerability, internal susceptibility and joint effects. This would, for example, allow for appropriate assessment of life stages, which is internal, from that of increased exposures near busy roads, which is external. is also the need to examine when both occur at the same time and joint effects are best able to explain risk. This comment also applies to Table 3-9 on page 3-29. There

Figure 2-7. It is important to include in the Evidence-based Considerations box that along with "newly identified at-risk populations" there may be more evidence to consider for those at-risk groups already identified due to their external vulnerability or internal susceptibility.

Pages 3-1 – 3-2. This paragraph describes the assessment of impact on selected health outcomes. I recommend that the ISA integrate the discussion of groups with external vulnerability or internal susceptibility (i.e., at-risk groups) into the main health assessment portions of the document. This makes the information easier to understand rather than to separate this discussion into a section detached from the other health outcomes.

Dr. Terry Gordon

1. This is an IRP document used many times previously in NAAQS development planning. As such, it has been written and reviewed by over 4 dozen individuals and is of exceptional quality and clarity.
2. This is a bit of a departure from the norm, but it is unclear why the IRP and particularly its review is still needed to start off the development of each NAAQS review. As it doesn't really change, it could just be designated as a Standard Operating Procedure for all NAAQS reviews. As Cote, Costa, and Wagner (and others) drove the much needed change from AQCDs to the current ISA/PA process, EPA should make other changes and improve the efficiency and timing of NAAQS by removing the IRP step. In particular, whereas Vol 1 addresses the background and history of each individual NAAQS under review and is necessary (included in the ISA), Vol 2 of the IRP is not necessary.
3. In a similar vein, the kickoff meeting and request for public and scientific comments seem like a waste of time. The EPA staff does a stellar review of the published literature that surely encompasses 99+% of all the information derived from the 'kickoff' step, so why not let public and scientific comments come after their review. Again, this would require a change in the 'process'.
4. Page 2.3, para 1 – This is a very important description of the 'process' used in answering the question above it. The footnote regarding the role of the PA in this process is also important and relevant to this paragraph and it should perhaps not be relegated as a mere footnote.
5. Page 3.9 and Section A.2 – Exposure Science and Dosimetry, while important in understanding health effects, surely don't seem to change much over a 5-year schedule, particularly for Dosimetry. The time devoted to it by EPA staff should/could be reduced – the question is whether atmospheric science and dosimetry are really useful in deriving a NAAQS (except maybe PM and particle size considerations)?
6. Table 3-7 – As one of the key elements of the 'process', why not move the excellent Causality description found at the end of the document into this section?
7. Page A-3 – In this IRP, do co-pollutants get sufficient 'time'? They are more than just contributors to the uncertainty discussion.
8. Page A-7 – The stated two orders of magnitude might be ok for tox studies for some criteria pollutants but not for some such as ozone, for example.
9. Minor Comments
    a. Page 2.2 – why include 'largely' in line 1, para 2? Seems a bit wishy-washy.
    b. Page 2.3, footnote 11 – Is this consistent with the definition provided on page A-52?
    c. Table 3-1 – I don't believe the division into short-term and long-term sections for the PECOS statement is necessary.

d. Page 3.8 – This new/different use of 'STEM' is a bit confusing given the broad use of the acronym in science.
e. Figure 3-1 - Does EPA take a random sampling of these excluded publications to see if they were correctly excluded?
f. Figure 3-2 – was the hepatic system actually studied in 2 controlled human exposure studies?
g. Page 3.22 - Some of the questions in this section are really similar to the chemistry ones above this section.
h. Page 3.23 – I would suggest deleting the question: "What NO2 and/or NO reaction products, including oxides of nitrogen metabolites, can be found in the cells, tissues, or fluids of the respiratory tract and in the systemic circulation that may serve as markers of NO2 and/or NO exposure and effect?".
i. Page 3.28 – I would suggest deleting this sentence because there is so much uncertainty in the data: "What information is available regarding the effect of long-term, low-concentration exposure to oxides of nitrogen on an individual's sensitivity to short-term but higher concentration exposures?"
j. Table 3-9 and the following section are excellent and clear.
k. Table A-2 – For the Comparison group, I think this may be too strict. For example, EPA's initial PM2.5 controlled human exposure studies used whatever ambient concentration and concentration factor they got that day to develop an important dose-response study.
l. Figure A-2 – delete CASAC and WoS definitions as they're not in the figure (at least I don't see them).
m. Figure A-3 – the diagonal lines in the Mortality column are unclear.
n. Page A-21, line 8 – should 'between' be 'among'?
o. Table A-7 – should 'particle size' be included in this generic table (obviously, not necessary if NOx only).
p. Section A.7.2.1 – this Causality section is so important that it might be moved up to the beginning of this document.
q. A-44, last paragraph – particularly clear
r. A-55 – why say 'generally chaired by a CASAC member'? Shouldn't it be an experienced CASAC member familiar with the process?

Dr. Michael T. Kleinman

1. The focus of the document is NO2, which is the indicator for the current standard. However the document mentions that other nitrogen compounds, which are relevant, should be discussed. The ISA should specifically address any circumstances where other chemical species might be relevant, or where NO2 concentrations might underestimate health effects and address the question on pg 2-5 of V2 "Does the currently available information call into question the use of NO2 as the indicator for the primary standards for oxides of nitrogen? Is support provided for considering a different indicator?".
2. Pg 3-3 Scope – For areas where research efforts have subsided EPA should consider whether new research is needed and should be stimulated to address gaps in knowledge. This might be relevant to the stipulation that AQ Criteria include consideration of "nitric and nitrous acids, nitrites, nitrates, nitrosamines, and other carcinogenic and potentially carcinogenic derivatives of oxides of nitrogen."
3. Pg 3-4, para 2, L2-4: the list of health outcomes should include possible neurological effects, for which there are some publications, see references below.
4. P3-18 Table 3-7: The criteria for causal and likely to be causal are quite restrictive. The weight given to controlled human studies is relevant but whether it is overvalued needs to be considered, especially if differences lead to lessening the overall causality description. Some limitations include the generally small numbers of individuals studied can reduce the power to detect significant changes, the individuals studied are not representative of the most sensitive individuals, the exposure usually does not represent some of the reactive nitrogen species that are present in ambient air.
5. Table 3-9 (pg 3-29) should be carefully re-evaluated. Since 2016 there were several papers that demonstrated significant cardiovascular effects of nitrogen oxides. Many of the studies were not performed in the US. Especially for studies in which the effect is measured in terms of changes induced by an incremental increase in pollutant exposure, whether or not the ambient concentration was < 22 ppb might not be a detriment.
6. Appendix A; The literature screening protocol is very well defined and commendable.

References

Freire, C., et al., *Association of traffic-related air pollution with cognitive development in children.* J Epidemiol Community Health, 2010. **64**(3): p. 223-8.

Petkus, A.J., et al., *Associations Between Air Pollution Exposure and Empirically Derived Profiles of Cognitive Performance in Older Women.* J Alzheimers Dis, 2021. **84**(4): p. 1691-1707.

Yan, W., et al., *Acute nitrogen dioxide inhalation induces mitochondrial dysfunction in rat brain.* Environ Res, 2015. **138**: p. 416-24.

Ye, S., et al., *Ambient NO(2) exposure induces migraine in rats: Evidence, mechanisms and interventions.* Sci Total Environ, 2022. **844**: p. 157102.

Ye, S., et al., *Mechanism of NO(2)-induced migraine in rats: The exploration of the role of miR-653-3p/IGF1 axis.* J Hazard Mater, 2024. **465**: p. 133362.

Dr. Michelle Oakes

The Integrated Review Plan (IRP) is overall a well-written and thoughtful document. The IRP clearly outlines the scope and organization of the Integrated Science Assessment (ISA) as well as the general review process for including and evaluating new and existing scientific information into the ISA by discipline. Below are a few comments/suggestions for clarifying topics within the document.

Comments/Suggestions:

1) In Section 2 "Policy-Relevant Issues in the Current Review", on page 2-4, the IRP states the following question: "Is there evidence of effects at oxides of nitrogen exposure concentrations lower than those at which were previously observed or in areas that would have likely met the current primary standards?" In many cases, NO_2 concentrations are expected to vary between near-source (e.g. near-road environments) vs. central site, urban-scale locations in the same metropolitan area. Will the NAAQS review process take these differences into consideration when determining whether an area is likely meeting the current primary standards?
2) The question, on page 2-4, "To what extent is new information available to improve our understanding of the NO_2 concentration gradients around important sources, such as major roads and combustion sources, and how those gradients relate to ambient air monitoring concentrations across larger areas?" is important. This question addresses spatial differences, but it does not appear to consider potential temporal differences that may provide unique context to exposure differences and background concentrations across larger areas, in both urban and remote settings. Addressing both spatial and temporal differences may help interpret health effects data by identifying potential exposure uncertainties.
3) In the Atmospheric Sciences section, the use of low-cost sensors for characterizing air quality & exposures, filling in spatial gaps and understanding co-pollutant trends is an area of significant research growth since the last NOx ISA. Further, low-cost sensor data maintained on EPA resources/tools, such as the Fire & Smoke Map, are being widely used, at least in the air quality management/regulatory space. In several cases, these sensors have been shown to provide reasonable data which can be used to supplement more high-quality data. However, data quality and sensor degradation rates vary widely across sensors. In the ISA, EPA may want to limit sensor studies to those that have performed rigorous performance evaluations against regulatory grade monitors and have

periodically conducted standard quality control/quality assurance measures. If unique information is provided in a sensor study lacking these performance and QA/QC procedures, the limitations need a thorough discussion, and any interpretation from these studies needs to be qualified with potential uncertainties.

4) In the *Source to Concentration –Air Quality, Atmospheric Science, Fate and Transport section*, satellite, low-cost sensors and mobile source monitoring data can be useful in characterizing spatial variations of NO_2 and co-pollutants, particularly around unique local sources and exceptional air quality events. These data sources can be particularly useful in a discussion of air quality at the near-source or neighborhood scale. Additionally, analysis of near-road data collected from monitoring networks operated by state and local air agencies will be useful in determining peak exposures and co-pollutant exposures during different seasons and meteorology conditions. This type of analysis will help address key uncertainties outlined in the 2016 ISA.
5) Regarding targeted air quality analyses in the footnote numbered 33, which type of data or data sources does EPA intend to use for targeted analyses?

Minor Comments:

1) Recommendations based on the NASEM committee are cited frequently throughout the Draft IRP Vol 2. Would it be useful to have table in the IRP (or the ISA) summarizing the NASEM key recommendations?
2) In Appendix A, Section A.1. The first paragraph appears to repeat a few sentences and probably needs editing.
3) In the Introduction, on page 1-1, the IRP states that "Consistent with the reviews completed in 2010 and 2018, this review focuses on health effects associated with gaseous oxides of nitrogen and the protection afforded by the primary NO_2 standards. The gaseous oxides of nitrogen include NO_2 and NO, as well as their gaseous reaction productions." This language should be consistent throughout the IRP across disciplines. It is worthwhile to check for consistencies throughout the document.

Dr. Richard E. Peltier

While this may be implicit in its wording in the IRP, I recommend EPA revise the requirement that work considered in the ISA is subjected to ethical review to include both Institutional Review Board oversight (as documented in this plan and in the attached Appendix A), but also by foreign equivalents of IRBs. Without an explicit statement, important peer-reviewed research conducted in other nations may be plausibly excluded.

Section 2.0: consider rephrasing "…current standards provide requisite protection of public health, and his decisions as to whether..." to "…current standards provide requisite protection of public health, and their decisions as to whether..."

The plan (section 3.2) notes the evidence for organic and inorganic nitrate risks as summarized in the 2019 ISA for PM, and recently reconsidered. While these are important particulate-bound forms of nitrate, relative to this review their presence is important in terms of particle/gas partitioning[1] and could pose an important source of NOx exposure where few primary emissions were expected.

I applaud the expanded inclusionary criteria described within the PECOS toolsets, and considers science based on quality and merit rather than geography – I believe this is the right choice. That said, this is a different inclusion framework than, for example, a recent CASAC ozone review, and EPA might consider including a short narrative that describes and justifies this expansion.

On a related note, it may be useful to provide a narrative in which EPA can describe how important prior work that may fall outside the temporal scope of this review ('definitive works in the literature', section 3.2) will be evaluated and included or excluded. If few/no updated studies have been produced relative to this review, what criteria does EPA use to identify these older 'definitive works'? Or in cases where important works are identified by some mechanism (public comments, staff notation, CASAC recommendations, etc), does EPA ignore the date restrictions and simply apply the PECOS or STEM rubrics?

Given that NOx is one of the few NAAQS with significant indoor sources (where exposures are likely to be important, even though regulatory concerns are limited), EPA may wish to consider how the impact of indoor exposures by certain populations – and not others - may lead to increased susceptibility to adverse health outcomes. While EPA rightly acknowledges the potential for confounding, they may wish to interrogate whether, for example, in a community

[1] Aruffo et al, EAS&T, 2022

where indoor gas cookstove use leads to increased indoor NOx concentrations[2] and exposures, does this lead to a population that is increasingly susceptible to adverse health outcomes under more modest ambient NOx conditions?

EPA intends to use a set of machine learning tools to maximize efficiency (SWIFT-Active Screener and Living Literature Review; and perhaps DistillerSR). This is certainly a practical and useful approach to data reduction, but it may benefit EPA to discuss potential limitations and risks to using such tools. I do not recommend avoiding the use of these important tools, but do think the public would benefit from a better understanding of the (likely low) risks of error in this critical data reduction step that better illustrates the rationale or reasoning for these exclusions.

[2] E.g. Zhao et al, Indoor Air 2020

Dr. David Rich

Overall, the Oxides of Nitrogen IRP Volume 2 was well written with adequate descriptions of key concepts that will be used as part of the ISA. The Appendix provide necessary detail on most sections and components. However, there were a few portions of the text for which I have questions, comments, or would suggest some revisions.

1. **PAGE 11 – "Standards to be set at a level that avoids unacceptable risk."**

   What defines an unacceptable risk? Acceptable to whom? It is not clear what is an acceptable versus an unacceptable risk. This likely has been used in past IRPs and ISAs, but it should be expanded on a bit here.

2. **PAGE 12 – "As in past reviews of the primary NO2 NAAQS, this will likely include a focus on people with pre-existing respiratory disease, children, and older adults." PAGE 46 – Discussion of methods to identify at-risk populations.**

   These are just two of the many places in the document discussing at-risk populations and their importance and use in the ISA. Overall this discussion is adequate. However, of all the methods given on how at-risk populations will be identified, why are past ISAs for other criteria pollutants not used as a starting point for the oxides of nitrogen ISA? If an at-risk population was identified in that ISA, I suggest it should be included on the list to consider for the oxides of nitrogen ISA. Also in Figure 2-1 – a question is asked if new at-risk populations were identified? However, what will be the metric or rule to decide if 1 or more "at-risk populations" have been identified across the published studies identified?

3. **PAGE 13 – "At what pollutant concentrations" and "…range of ambient air concentrations within which oxides of nitrogen contribute to health effects."**

   Some text should be included to define what statistical measures provided in published papers will be used to define the concentrations of oxides of nitrogen of that study. It seems likely that its just a mean or median or a range, since those are most often reported, but please make it clear.

4. **PAGE 13 – "To what extent are health effects found to be associated with oxides of nitrogen in epidemiologic studies being elicited by oxides of nitrogen exposure versus exposure to one or more co-occurring pollutants (e.g., PM2.5, CO, O3, SO2, other traffic-related pollutants)? "**

How will EPA make this judgement? Is this just seeing if multiple pollutants perhaps thought to represent traffic pollution (e.g., NO_2, CO, black carbon) are all associated with the health outcome in the study in separate one pollutant analyses, or will more emphasis be given to studies that conduct 2 or more pollutant analyses (i.e., assessing confounding), or even those that also assess mixture effects?

5. **PAGE 55 – section A3 – "Studies examining higher exposure concentrations (i.e., one to two orders of magnitude greater than ambient air concentrations) may be included if they provide evidence of the potential biological mechanism(s) for an observed effect."**

   This text is not addressing the specific issue I am raising, but are studies done outside the US to be included in the oxides of nitrogen ISA? And if so, is it only in this context, i.e., to provide evidence of biological mechanisms? Will studies in Canada be considered, because of proximity to the US, if studies from other countries in Europe, Asia, and elsewhere will not? Such a statement providing the answer to this question should be repeated several times in the document.

6. **PAGE 73 – "Biomarkers: For some pollutants, epidemiologic studies use biomarkers to estimate exposures. As noted above for personal monitoring, biomarkers provide exposure estimates at the individual level and that are attributable to both ambient air and non-ambient-air sources. Depending on the biomarker used, exposure biomarkers may be quite limited with regard to the specific timing and duration of the exposure represented. When used, biomarkers should be clearly justified and measured using valid, reliable methods with appropriate characterization of variability."**

   Biomarkers of NO_2 or other pollutant exposures can be useful for evaluating the association between traffic pollution or NO_2 exposure and a health effect directly. However, how that biomarker values corresponds to a specific ambient NO_2 concentration at a monitoring site with minimal uncertainty to allow it to be given weight in determining if a specific NO_2 standard can protect the population may be questioned. Can an example of one such past use of such a biomarker study be included in the text to document how this has been done in past ISAs?

7. **PAGE 74 – "A confounder is associated with both the exposure and the outcome. Factors are considered to be potential confounders if demonstrated in the scientific literature to be associated with both the exposure and outcome being evaluated."**

A minor point, as the discussion of confounding below this section is appropriate and detailed. Specifically, Section A.5.1.7. and A.5.1.8 - Discussion of confounding and Statistical methodology are well written. However, this definition of confounding on page 74 should be corrected. The text should state that variables are considered to be a confounder if they are correlated with the exposure, are risk factors for or predictors of the outcome independent of the exposure, and are not on the causal pathway from exposure to disease. This should be changed so that variables that are products of exposure are not considered confounders and thus lead to incorrect analyses and inference.

8. **PAGE 43 – "What evidence is available regarding the nature of health effects from exposures to ambient air pollutant mixtures that include oxides of nitrogen? To what extent does the evidence support attributing these health effects to exposures to NO2 or other oxides of nitrogen, another ambient air pollutant that is correlated with oxides of nitrogen, or to the pollutant mixtures that oxides of nitrogen may be representing?"**

   How will the EPA decide whether the evidence supports attribution of health effects to oxides of nitrogen, another pollutant, or to the mixture itself? Will 2 pollutant models to control for co-pollutant confounding be used in the same way as in past ISAs to address this? The question asks what evidence is available, but does not say what statistical analyses or pieces of information from published studies will be used to make such determinations. An example could be given to clarify this.

9. **PAGE 42 – "What do recent studies indicate regarding the health impacts of reductions in concentrations of oxides of nitrogen in ambient air (e.g., due to policy intervention) or reductions in exposures (e.g., due to changes to indoor sources)? "**

   This reads as though indoor sources are components of the oxides of nitrogen for which EPA is assessing associated health effects, when I assume that is not part of the regulated oxides of nitrogen. If not, can this be modified to make that clear?

10. **PAGE 43 – BOTTOM – "To what extent does recent data from epidemiologic, controlled human exposure, and animal toxicological studies provide information on health effects related to various short-term exposure durations (e.g., 1-hour, 24-hour, multi-day)?"**

    I assume these are not the only durations of air pollution exposures to consider when deciding if there are short-term associations between a health outcome and oxides of nitrogen. Will durations of a few hours (e.g., 3-hour average, 6-hour average, 12-hour average) also be given equal weight in deciding if short term exposures are associated

with a health outcome, or will those with 1 hour or 24-hour durations be weighted more heavily?

11. **PAGE 44 – "Do recent studies provide information on health effects related to long-term exposure windows other than annual or lifetime average (e.g., preconception, pregnancy average, pregnancy trimester average)? What data are available comparing associations of health effects among various long-term oxides of nitrogen exposure metrics (e.g., annual, seasonal, pregnancy average)?"**

    Are all pregnancy exposures considered long term, or if we are looking at weekly exposures during pregnancy and perhaps a several consecutive weeks (as is done when using distributed lag models with gestational week pollutant exposures in the model), are these long-term exposures or short-term exposures? Is a trimester average considered long term? For studies of pregnancy health effects, this is an important consideration given the expanded use of distributed lag models in such studies. In future ISA's, these analytic models will be even more frequent in the published literature on air pollution exposures during pregnancy and maternal and fetal health effects.

<u>Minor correction</u>

12. **PAGE 81 – Section A.5.2.7 – last sentence – "Consistent trends across studies can informative, even if results of individual studies are not statistically significant."**

    Add "be" between "can informative"

Dr. Lianne Sheppard

IRP Volume 2

General comments

1. Overall, the IRP Volume 2 is a helpful document that provides a good foundation for the ISA review. It will benefit from revision based on this consultation with the CASAC.
2. The document should clarify how the Volume 2 Appendix A is different from the ISA Preamble and articulate the planned next steps regarding the Preamble.
3. There should be more explicit text about the changes that were made as a result of the NASEM advice or the IRP should refer to the document where this information can be found.
4. I think it would be helpful to incorporate into Chapter 2 a discussion of the previous CASAC advice provided in previous review(s), EPA's decision in response to that advice, and the EPA's plans to address that previous advice, if any, in the current review. Particularly, if there was any discrepancy between the CASAC recommendations and the EPA final decision in the past review, or the CASAC provided recommendations for future reviews, these should be documented in the IRP and a plan for how they will be followed up on should be stated.
5. I'm surprised that Chapter 3 does not include a list of anticipated health effects chapters. I think this should be added. At least partially this list can be inferred from Table 3-8, but this seems inadequate.

Introduction – Chapter 1

P 1-1 "conversation" or "conversion"?

P 1-1: It would be helpful to indicate that the "general approach for this review" is documented in Chapter 2 and the planning considerations for the ISA development are in Chapter 3, particularly since the title of Chapter 2 is not "general approach for this review".

P 1-2: Volume 1 lists milestones, but it omits including a schedule. Is this intentional? When will CASAC and the public be informed of the planned schedule?

Policy-Relevant Issues in the Current Review – Chapter 2

P 2-1: The text states that the second overarching question will only be considered "as appropriate". While on the surface this approach seems reasonable, in practice it has been problematic. I believe that we want to guard against this "as appropriate" framing because it can result in a too restrictive set of possibilities considered and presented to the CASAC and the Administrator. As an example, in the recent ozone review, the second question regarding alternative standards was not addressed with any depth because the EPA staff concluded that the available scientific information did not call into question the adequacy of the protection afforded by the current standard(s). This conclusion was reached by the EPA staff in spite of the presence of multiple previous recommendations by the CASAC over more than a decade, along with multiple public comments, arguing that the standard should be lower. This topic will be addressed further in the April 25 meeting of the Chartered CASAC regarding the NAAQS review process.

P 2-3 first paragraph: I think this paragraph should also acknowledge previous advice from the CASAC and the public. Also, this input should be considered explicitly in addressing the question at the top of this page, rather than ignored or handled separately.

P 2-5 paragraph following the second bullet: Again, this text does not adequately acknowledge the conclusions and advice from previous reviews.

P 2-5 last bullet: Consideration of the available analyses regarding exposures and risks could be problematic if those analyses are insufficiently comprehensive. Again, if previous CASAC advice is not explicitly incorporated into the review, the available analyses may not address the full range of alternative standards that the CASAC recommended considering in past reviews.

Figure 2-1: Note that the Chartered CASAC is recommending that this flow chart and decision process be revised, as will be discussed at the upcoming April 25 NAAQS process meeting. This figure inherently restricts the kinds of alternative standards that are considered in the case that the EPA staff conclude that the information does not call into question the adequacy of the current standards, even when the CASAC (or the public) may have provided different advice in the past.

Development of the ISA – Chapter 3

P 3-2, last paragraph, text beginning "The approach described in Appendix A…": Does this mean that the ISA Preamble will be updated to reflect recent advances and the NASEM advice? If the evolving process is making the Preamble out of date, then I think the Preamble should be

updated. (Note that the Preamble was published in 2015.). This comment is also relevant to text on p. 3-10.

Table 3-1:

- Comparison: Per unit increase implies a linear dose-response. Is this the only comparison to be considered or should the comparison be restated? The latter part of this cell covers possible nonlinearity, though the "e.g." could be expanded to other approaches to modeling.
- Outcome: Any reason to expand beyond incidence and prevalence? For instance, how will lung function decrements be handled?
- Study design: The listed designs don't seem to be specific to the short-term and long-term exposure categories. When will a cohort study design be used to assess short-term exposures? When will panel, case-crossover, or time series studies be used to assess long-term exposures?

Table 3-2:

- Outcome: I think some of the effects of interest listed will apply to some epi studies as well.

P 3-10: There is no mention of the workshop that EPA typically holds. How does that input fit into the process described in Section 3.3?

Literature flow (e.g., Figure 3-1): Is EPA harnessing AI for any of the steps in the process, at least as a preliminary processing step? If so, I think it should be documented. There is reference to "machine ranking tools" (p 3-12), which may be an indirect reference to AI.

P 3-15, last line: I suggest the "focus on validated models used to estimate exposures" is worth some discussion. What are the model validation criteria that will be considered?

P 3-16 strength of study design and footnote 24: I am concerned that the outside North America exclusion may be too restrictive. How might an outside North America study still be useful yet not meet the stated criteria? It would be useful to see some examples of how this will be applied. Furthermore, given that studies outside of North America can be eliminated in the preliminary evaluation, should there be further consideration of this exclusion criterion?

P 3-17: I appreciate the referencing of and attention to the Savitz et al papers.

P 3-17 last sentence: The EPA should articulate a plan for updating the Preamble.

P 3-19 end of Section 3.3: I agree that one ISA draft is ideal and should be the goal. However, I am concerned that there is no provision in the IRP for a second draft ISA should the CASAC feel that this is needed. I think the text should be modified to leave open the possibility of a second draft ISA if the CASAC provides this advice.

P 3-21 exposure subsection: Another question to add: Are there new insights from exposure assessment study design that have implications for inference about health effects? This could be folded into the third bullet about exposure measurement error on p. 3-22.

Appendix A: ISA Development Process

Please clarify the specificity of this appendix: How generic is it to all NAAQS pollutants versus specific to NOx?

Section A.1: Text in this section is repeated verbatim multiple times.

P A-21 text on short-term exposure studies: I think it is more appropriate to lump case-crossover studies with time series studies rather than with panel studies. Case-crossover studies are like time series studies except that they use matching rather than modeling to control for time-varying confounders. Panel studies are typically much smaller than case-crossover studies and are designed to address different questions. For instance, there are no panel studies of a subject such as out of hospital cardiac arrest, which this could be considered in a case-crossover study. I think the prioritization of panel and case-crossover studies over time series studies is problematic.

P A-21-22 text on long-term exposure studies: Why is attention to cohort studies restricted to prospective cohort studies?

P A-42: Where in the process does the peer input workshop happen? I note that this is not included in the milestones table in Volume 1.

<u>IRP Volume 1</u>

Preface, p 2-2 last complete paragraph, and Table 4-1: The reference to Volume 3 in the Preface states that the risk and exposure analyses will be included in the PA; Table 4-1 makes the same statement. The second sentence in the p 2-2 last complete paragraph implies this. In other reviews, the CASAC has had oral discussions with the EPA staff about their decision to no longer provide a stand-alone REA and I think the EPA staff's rationale should be documented in writing. I recommend that some text providing EPA's rationale for this approach be added to Volume 1.

P 1-3 first paragraph: The text notes that the CASAC "shall complete a review of the criteria". The EPA has not asked the CASAC to the review the list of criteria air pollutants and, given historical practice, the CASAC, unlike the SAB, typically only responds to requests from EPA. The current six criteria air pollutants are not listed in the Clean Air Act (CAA). The CAA only specifies criteria air pollutants in general terms as those pollutants the "emissions of which, in his judgment, cause or contribute to air pollution which may reasonably be anticipated to endanger public health or welfare"; "the presence of which in the ambient air results from numerous or diverse mobile or stationary sources"; and for which he "plans to issue air quality criteria…". The air quality criteria are also intended to "accurately reflect the latest scientific knowledge…". Given that the current list of six criteria air pollutants has not changed in fifty years while scientific knowledge has advanced considerably, I believe that a scientific review of the criteria air pollutants is warranted. This review would include consideration of the current criteria air pollutants and whether there are pollutants that should be added to or dropped from the list. This is an important CASAC responsibility that should occur in the near future.

P 2-6 last two sentences: There is no mention of the possibility of a second draft PA.

Table 4-1 and related text:

- The milestones do not include any schedule. Presumably this is because of the unresolved citizen suit. However, I think this document is incomplete without mention of a tentative schedule for the review.
- There is no mention of any workshops. Shouldn't these be included?
- There is no provision for a second draft of either the ISA or the PA. While I applaud the EPA staff for their efforts to create high quality first draft documents, there should be provision for second drafts included as part of the milestones, specifically when second drafts are recommended by the CASAC.

Dr. Neeta Thakur

Section 1

-Page 1-1 states "Health effects and non-ecological welfare effects associated with the particulate species are addressed in the review of the NAAQS for particulate matter (PM). The EPA is separately reviewing the ecological welfare effects associated with and the secondary standards for oxides of nitrogen, oxides of sulfur, and PM." Does this mean these topics will not be included in the NOx ISA? If this is the case, for clarity re: the scope of the document, it would be helpful to explicitly state this.

Section 2

-In section 2.1, page 2-2, paragraph 2 there is mention of "sensitive groups" without a definition. In the same section, page 2-3, paragraph 1 there is mention of "at-risk populations", which is defined to include "people with pre-existing respiratory diseases, children, and older adults". Are "sensitive groups" and "at-risk groups" one and the same? Recommend using one term consistently, defining early in the ISA, and including how the definition was determined.

-Section 2.1 refers to "at-risk groups" and specifically includes those with respiratory conditions. Further in this section, the questions to be explored, include questions about looking at this group specifically concerning exposures at low levels of NOx. This is appropriate given the body of literature connecting NOx to chronic respiratory conditions. In Section 3.2.1, there is a list of the other health conditions that will be included as part of the review. If the evidence is high for one of the other conditions, would this health condition be considered in the "at-risk" assessment, or do these need to be pre-specified for the upcoming PA?

-The background documents (Volume 1) suggest that there were limited human exposure studies examining health effects at levels approaching 100ppb and lower (specifically between 100 – 250 ppb), in epidemiology studies there is a limited measurement of acute exposures during reported study periods (rather events are assumed to be averaged over the entire period) and concluded that the lack of associations with asthma visits in areas meeting annual standards were unlikely to exceed short-term exposure standards. In Volume 2, section 2.1, where the questions to be explored in the ISA are listed, I recommend adding a question to examine short-term exposure to NOx on health and across at-risk groups. While there may not be a sufficient number of studies, this is an important area to explore since it was named as a limitation in the previous assessment.

Section 3

-Section 3.2.1 states results will be integrated into the 2016 ISA. Will studies previously included in the ISA be re-evaluated? If not, have the review criteria for study inclusion changed in any meaningful way from 2016? What about studies that were included in pre-2016 ISAs, has the review process changed?

-Table 3-1. Exposure is limited to studies that are from the U.S. What is the rationale for not including all North American studies?

-Questions outlined in sections 3.4.2 through 3.4.6 are comprehensive and appropriate. Appreciate the inclusion of assessing data quality and modeling from other data sources (i.e., no EPA monitors) in the atmospheric section and the separation of questions assessing short and long-term health outcomes.

-Section 3.4.6 discusses how “at-risk groups” are defined. What is the difference in evidence stating there is “suggestive evidence” vs “adequate evidence” for an “at-risk” group? Will both levels be included in the ”at-risk” assessments for the ISA and PA?

Dr. Corwin Zigler

The Draft of the Integrated Review Plan for the Primary NAAQS for Oxides of Nitrogen, Volume 2 ("Draft IRP") provides updates to the description of the Integrated Science Assessment (ISA) causality framework. These updates make effective use of the recent National Academies of Sciences Engineering and Medicine report on Advancing the Framework for Assessing Causality of Health and Welfare Effects to Inform NAAQS Reviews ("NASEM Report"). These updates serve to a) clarify and reiterate the role of the ISA causality determinations and b) specify how individual studies employing causal inference/analysis/modeling methods will be evaluated for the ISA. I believe this is a positive evolution of the description of the framework for causality determinations, and that these updates will help to resolve apparent confusion in recent NAAQS considerations surrounding the scientific rigor of the causality determinations and the appropriate way to evaluate studies framed as "causal analysis" or "causal inference" in the weight of evidence. Most of my comments relate to possible refinements of the updated description of the ISA causality framework.

1. The Draft IRP makes clear that the description of the ISA causality framework has been updated in light of the NASEM Report. However, I did not find any high-level description of *how* the description of the causality framework was updated. Since the fundamental weight-of-evidence considerations, causal determinations, and justification of the ISA causality framework have *not* changed, it may be helpful to provide some clear and accessible sign posting of the updates to avoid the misconception that the overall framework is different from previous ISAs. The fact that the ISA causality framework has not fundamentally changed in light of the NASEM Report seems an important point that should not be lost.
2. I did not notice any explicit reference to the recommendations from the NASEM Report. Explicitly linking updates to the ISA causality framework that were made in response to recommendations from the NASEM may serve to bolster the justification for the updated framework.
3. The Draft IRP, particularly in Section A.5.1 ("Epidemiology"), provides needed focus on how to evaluate individual studies employing causal inference in the weight of evidence, making the essential point that the process for evaluating the validity of such studies is consistent with the process laid out for more traditional regression analysis approaches.
4. The Draft IRP settles on the term "alternative methods for confounder control" to describe certain causal inference methods while avoiding confusion with the ISA "causality determinations." The word "alternative" seems to be in relation to more traditional regression modeling. I fully support a labeling of methodologies to fend off any confusion between analysis approaches used in individual studies and the overall

causality determinations. However, I have concerns about the term "alternative methods for confounder control" as a moniker for causal inference methods because confounding is the primary, but not the only threat of validity targeted by this class of methods. Having said that, I struggle myself to come up with a more effective term, but would be curious if EPA considered alternative terminologies or if other CASAC members might have suggestions. In the absence of a clearly superior alternative, the current strategy of the Draft IRP to explicitly clarify that the term "alternative methods for confounder control" is used to describe a class of methods often referred to in the literature as causal inference or causal modeling is effective.

5. Section A.5.1 has repeated reference to various methods or study designs in terms of their intention to mimic randomized experiments towards support of causal validity. This is effective framing of a wide range of alternative methods for confounder control as well as quasi-experimental study designs. Along these same lines, there is an increasing trend in clinical medicine and epidemiology towards "trial emulation approaches" [Hernán and Robins, 2016]. These approaches are not distinct from what is currently described in the Draft IRP. I view them more as a popular and useful packaging of many elements of causal inference methodology. While trial emulation has mostly been used to frame comparative effectiveness studies of clinical therapies, the increasing popularity of this framing may lead to its extension to studies relevant to future ISAs. EPA may wish to allude to these approaches explicitly alongside the other methodologies and study designs noted for their attempt to mimic randomized trials.
6. Page A-27 describes approaches to handling unknown confounders. There is an opportunity to link these ideas to previous discussions of study designs (e.g., case-crossovers or quasi-experiments in A.5.1.1) as one key feature of these study designs is their potential to adjust for unknown or unmeasured confounders. In addition, it may be worth noting here some approaches common to econometrics (e.g., fixed effects, first-differences, difference-in-differences) that are similarly motivated [Greenstone and Gayer, 2009].
7. Page A-21 states that "Across study designs, studies with larger sample sizes and those conducted over longer time periods reduce selection bias among the study population and increase generalizability, and such studies are therefore considered to produce more reliable results than studies with smaller sample sizes." I do not agree with this statement as written, and would not equate the benefits of large sample size to matters of selection bias, study time period, and generalizability of results. For a given study design, large sample sizes can increase statistical precision, which is a benefit relative to smaller sample sizes. But there is no guarantee that larger sample sizes reduce selection bias or generalizability, and in fact there is recent literature articulating this point, for example, Meng [2018], Bradley et al. [2021]. Larger studies have *potential* to reach larger swaths of populations or longer time frames, and as such *may* present benefits in relation to generalizability or selection bias, but this potential would have to be evaluated in context. Poorly designed large studies could perpetuate selection biases. The salient point that

more diverse or broadly represented study populations are more desirable to avoid selection bias and improve generalizability should be stated more clearly, without claiming that studies with larger sample sizes produce more reliable results.

8. Page A-23 of Section A.5.1.3 specifies that studies only reporting associations with undefined mixtures or their surrogates (e.g. distance to roadway) are not used to inform ISA conclusions. I find this statement at odds with some of the text in Section 3.4.6 ("At-Risk Lifestages and Populations"), where proximity to important sources of NOx emissions, and roads in particular, is cited as a possible determinant of increased risk. My interpretation is that the ISA will view proximity to sources (e.g., roads) as a viable feature to define at-risk populations, but not a viable proxy for exposure. Some clarification may be needed, as this would still seem to permit measures like distance to roadway to inform ISA conclusions.

Minor Comments

- Pages A-1 and A-2 appear to literally repeat several lines of text.
- I believe the reference to "general propensity scores" on page A-22 should be to "generalized propensity scores" [Hirano and Imbens, 2004, Imbens, 2000].
- Section A.5.2.6 on potential confounding in controlled human exposure studies contains text that is very similar to that used to describe epidemiological study designs. To me, this underscores the importance of viewing observational studies according to how well they approximate (randomized) controlled studies. It may be useful to note this in the Draft IRP.

References

Valerie C. Bradley, Shiro Kuriwaki, Michael Isakov, Dino Sejdinovic, Xiao-Li Meng, and Seth Flax-man. Unrepresentative big surveys significantly overestimated US vaccine uptake. *Nature*, 600 (7890):695–700, December 2021. ISSN 1476-4687. doi: 10.1038/s41586-021-04198-4. URL https://www.nature.com/articles/s41586-021-04198-4. Publisher: Nature Publish- ing Group.

Michael Greenstone and Ted Gayer. Quasi-experimental and experimental approaches to environmental economics. *Journal of Environmental Economics and Management*, 57 (1):21–44, January 2009. ISSN 0095-0696. doi: 10.1016/j.jeem.2008.02.004. URL http://www.sciencedirect.com/science/article/B6WJ6-4TN0KS5-1/2/5b124b364bf72c

Miguel A. Herna´n and James M. Robins. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. *American Journal of Epidemiology*, 183(8):758–764, April 2016. ISSN 0002- 9262. doi: 10.1093/aje/kwv254. URL https://doi.org/10.1093/aje/kwv254.

Keisuke Hirano and Guido W. Imbens. The Propensity Score with Continuous Treatments. In *Ap- plied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives*, pages 73–84. John Wiley & Sons, Ltd, 2004. ISBN 978-0-470-09045-9. doi: 10.1002/0470090456.ch7. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/0470090456.ch7.

G. W. Imbens. The role of the propensity score in estimating dose-response functions. *Biometrika*, 87(3):706–710, September 2000. ISSN 0006-3444. doi: 10.1093/biomet/87.3.706. URL https://academic.oup.com/biomet/article/87/3/706/293734.

Xiao-Li Meng. Statistical Paradises and Paradoxes in Big Data (i): Law of Large Populations, Big Data Paradox, and the 2016 Us Presidential Election. *The Annals of Applied Statistics*, 12(2):685–726, 2018. ISSN 1932-6157. URL https://www.jstor.org/stable/26542550. Publisher: Institute of Mathematical Statistics.