Clean Air Scientific Advisory Committee (CASAC)

A Federal Advisory Committee to the U.S. Environmental Protection Agency

April 24, 2024

EPA-CASAC-24-001

The Honorable Michael S. Regan
Administrator
U.S. Environmental Protection Agency
1200 Pennsylvania Avenue, N.W.
Washington, D.C. 20460

Subject: Consultation on the 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 (March 2024)*

Dear Administrator Regan:

The Clean Air Scientific Advisory Committee (CASAC) NOx Review Panel, hereafter referred to as the Panel, met on April 16, 2024, to provide a consultation on the 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 (March 2024)*.

The Science Advisory Board Staff Office developed the consultation as a mechanism to provide individual expert comments to the EPA for consideration early in the implementation of a project or action. A consultation is conducted under the normal requirements of the Federal Advisory Committee Act (FACA) (5 U.S.C. § 10), which includes advance notice of the public meeting in the Federal Register and opportunity for public input.

No consensus report is provided to the EPA because no consensus advice is given. Individual comments from the Panel are provided in Enclosure A.

We thank the EPA for the opportunity to provide advice early in the review process. The Committee does not expect a formal response from the EPA.

Sincerely,

/s/

Dr. Elizabeth A. (Lianne) Sheppard, Chair
Clean Air Scientific Advisory Committee

Enclosure

NOTICE

The Clean Air Scientific Advisory Committee (CASAC) is a chartered federal advisory committee, operating under the Federal Advisory Committee Act (FACA; 5 U.S.C. § 10). The committee provides advice to the Administrator of the U.S. Environmental Protection Agency on the scientific and technical bases of the National Ambient Air Quality Standards. The findings and recommendations of the committee do not represent the views of the Agency, and this document does not represent information approved or disseminated by EPA. The CASAC reports are posted on the EPA website at: https://casac.epa.gov.

U.S. Environmental Protection Agency
Clean Air Scientific Advisory Committee
NOx Panel

CHAIR
Dr. Elizabeth A. (Lianne) Sheppard, Rohm and Haas Professor in Public Health Sciences and Professor, Department of Environmental & Occupational Health Sciences and Department of Biostatistics, Hans Rosling Center for Population Health, University of Washington, Seattle, WA

MEMBERS
Mr. George A. Allen, Chief Scientist, Northeast States for Coordinated Air Use Management (NESCAUM), Boston, MA

Mr. Ed Avol, Professor Emeritus, Population & Public Health Sciences, Keck School of Medicine, University of Southern California (USC), Los Angeles, CA

Dr. Howard H. Chang, Professor, Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University, Atlanta, GA

Dr. Judith C. Chow, Nazir and Mary Ansari Chair in Entrepreneurialism and Science and Research Professor, Division of Atmospheric Sciences, Desert Research Institute, Reno, NV

Dr. Deborah Cory-Slechta, Professor, Environmental Medicine, University of Rochester Medical Center, University of Rochester Medical Center, Rochester, NY

Mr. Henry (Dirk) Felton, Research Scientist, Division of Air Resources, Bureau of Air Quality Surveillance, New York State Department of Environmental Conservation, Albany, NY

Dr. Mark W. Frampton, Professor Emeritus of Medicine, Pulmonary and Critical Care, University of Rochester Medical Center, Rochester, NY

Dr. Christina H. Fuller, Associate Professor, School of Environmental, Civil, Agricultural and Mechanical (ECAM) Engineering, University of Georgia College of Engineering, Athens, GA

Dr. Terry Gordon, Professor, Medicine, New York University School of Medicine, New York, NY

Dr. Michael T. Kleinman, Adjunct Professor, Environmental and Occupational Health, University of California, Irvine, University of California, Irvine, Irvine, CA

Dr. Petros Koutrakis, Director of Exposure, Epidemiology and Risk Program, Environmental Health, Harvard T.H. Chan School of Public Health, Boston, MA

Dr. Julian Marshall, Professor, Civil & Environmental Engineering, University of Washington, Seattle, WA

Dr. Michelle Oakes, Manager, Division of Air Pollution Control, Tennessee Department of Environment & Conservation, Nashville, TN

Dr. Jennifer Peel, Professor and Section Head of Epidemiology, Department of Environmental and Radiological Health Sciences, Colorado State University, Fort Collins, CO

Dr Richard Peltier, Professor, Environmental Health Sciences, University of Massachusetts, University of Massachusetts, Amherst, MA

Dr. David Q. Rich, Professor, Public Health Sciences, Medicine, and Environmental Medicine, University of Rochester School of Medicine and Dentistry, Rochester, NY

Dr. Neeta Thakur, Associate Professor in Residence, School of Medicine, University of California at San Francisco, San Francisco, CA

Dr. Corwin (Cory) Zigler, Professor of Statistics and Data Sciences, Department of Statistics and Data Sciences, University of Texas at Austin, Austin, TX

SCIENCE ADVISORY BOARD STAFF
Mr. Aaron Yeow, Designated Federal Officer, U.S. Environmental Protection Agency, Science Advisory Board Staff Office, Washington, DC

Enclosure A

Individual 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 (March 2024).*

Mr. George Allen

Overall, the IRP looks good – well written and complete.

General comments: the review completed in 2010 begat the near-road NO2 network of 79 sites. Those data show essentially no violations of the 1-h NAAQS of 100 ppb, and only rare exceedances. There is value in near-road measurements, but not for use in a NAAQS compliance context with existing daily or hourly standards for NO2, PM2.5 and CO. The existing infrastructure of the near-road network could be leveraged to support more detailed non-NAAQS measurements at a subset of the existing network at sites with high AADT counts and a high proportion of HDD diesel vehicles (LA I-405 for example). Hourly XRF, particle number concentration, BC, LDSA, oxidative potential, robust PM-coarse measurements are some examples. The near-road network is mentioned in passing on pg. 3-8 and 3-22 (TRAP bullet), but it may be useful for the ISA to explore this idea in the monitoring needs section.

This IRP notes the use of controlled human exposure studies. I would caution against relying on these as the primary source of evidence to the exclusion of epi studies; this was a major problem with last year's O3 reconsideration review.

Vol 1 summarizes the general approach to CASAC reviews. There are some process concerns based on recent reviews for O3 and the NOx-SOx secondary NAAQS re: EPA's stated objective of only doing one draft of a review document. In both of these cases, CASAC was not satisfied with the quality of the documents. Historically (10+ years ago), EPA staff would do revisions as requested by CASAC (examples: 2012 O3 ISA and 2013 PA) until CASAC said they were adequate for the intended purpose. Is that still the case, or are we really constrained to a single draft of major review documents? Slide 15 of the EPA presentation for this meeting presents milestones for this review, and only shows a single draft for each document. The slide notes a deadline suit has been filed, and that "EPA anticipates that resolution of those claims would inform the schedule for completion of the review." Having legal deadlines drive the quality of the review process would be an unfortunate outcome (see 2005 comments by McClellan, EPA-SAB-CASAC-05-012). EPA has gone back to the courts to request more time to comply with deadlines (e.g. multiple extension requests for the PM reconsideration final action in 2023 and early 2024); asking the court for an extension to allow revisions requested by CASAC to produce a better document would seem to be a reasonable request. Given that "deadline suits" have become the norm for NAAQS reviews (except CO so far), this is an issue that needs to be addressed.

Minor Comments

Pg. 2-4: bullet on confounding / co-pollutants. Epi studies of NO2 are often confounded by diesel PM (DPM) co-emitted with tailpipe NO2. Diesel is a major urban source of NO2, but health effects of DPM may contribute to or dominate the observed NO2 effects. Many studies, especially from Europe or other areas where automotive diesel is more common than in North America, focus on NO2 but do not control for DPM exposures. BC or EC can be a surrogate for DPM, but that relationship may be changing as the diesel fleet cleans up.

Pg. 3-1: I support the change from appendices to chapters in the ISA. I also support the Integrated Synthesis (IS) approach with details in the chapters. For most end-users of the ISA, the Executive Summary and IS is what will be used, with reference to the chapters as needed for the details.

Pg. 3-3, 2nd paragraph: It is important to carry forward older studies that provide important observations that are not part of more current literature; this was not always done in the 2020 Ozone ISA. An example is the Korrick et. al 1998 hiker study which was totally dropped from that review and was a significant factor in the O3 panel's discussions on the adequacy of the current standard.

Pg. 3-4, PECOS geographic limits. As noted previously I am somewhat concerned about constraining study inclusion to strict geographic limits such as North America. The footnote 24 on pg. 3-16 mentions different inclusion criteria for studies with populations outside of North America. Many studies, especially for NO2, come from areas like Europe that have higher use of automotive diesel.

Pg. 3-11: Do we know what the expected cut-off date for this ISA might be? June 2023 is noted here; is that the final cut-off date?

Pg. 3-15: Evaluation of study quality, also Section A.5 (pg. A-18). This is a critical component of study selection criteria that was an issue in the O3 reconsideration. Section A.5 notes "Individual study quality is evaluated by considering the design, methods, and documentation of each study, but not study results." To clarify my comment during the meeting, study results can be important when they imply highly implausible associations, and while a study should never be excluded solely on results, improbable results that are sharply distinct from the majority of the literature should trigger more careful review of the study's methods. The 2020 O3 ISA had three examples of this for long-term mortality that showed that O3 was protective (HR < 1.0) as shown in this excerpt from Figure 6-8 on page 6-29 of the 2020 O3 ISA:

†Eckel et al. 2016	CA Cancer Cohort	1988-2011	40.2
†Crouse et al. 2015	CanCHEC	1991-2006	39.6
†Cakmak et al. 2018	CanCHEC	1991-2011	15.0-43.0
†Weichenthal et al. 2017	CanCHEC	1991-2011	38.29
†Bentayeb et al. 2015	Gazel	1989-2013	40.5
†Carey et al. 2013	English Medical Practice	2003-2007	25.85
†Kim et al. 2017	NHIS-NSC	2007-2013	19.93

Hazard Ratio (95% CI)

There was no discussion of these implausible results in the ISA. Dr. Sheppard pointed out methodology issues with Kim et al. My individual comments for the March 2023 CASAC O3 PA meeting provide details on Bentayeb and Carey et al. studies and why they should have been excluded based on methodology.

Hvidtfeldt et al. (2019, https://doi.org/10.1016/j.envint.2018.12.010) also saw a HR < 1 when O3 was negatively correlated with all other pollutants and notes "Since a protective effect of O3 on mortality seems little plausible, we consider the lower HRs in association with higher exposure to O3 in our study to be a reflection of the inverse correlation between O3 and truly harmful pollutants." This is why I say that there can be occasions when study results should not be totally ignored when assessing study quality.

Pg. 3-20: bullet on new measurement methods / remote sensing. The TEMPO satellite (https://tempo.si.edu/) will be operational later this year and provides very detailed hourly NO2 data for the entire US. This is a major improvement in spatial coverage and should be noted in this part of the ISA.

Pg. 3-22: bullet on multi-pollutant (traffic-related) co-exposures. This is a very important topic. Confounding by components of TRAP can make assessing NO2 health effects challenging. See above re: NO2 and DPM.

Mr. Ed Avol

IRP Volume 2 ("Planning for the Review and the Integrated Science Assessment") describes 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 for staff consideration:

P2-3, end of 1st paragraph, "As in past reviews, … this will likely include a focus on people with pre-existing respiratory disease, children, and older adults.'- – it is entirely appropriate to continue this approach with regard to susceptible sub-groups, but there may now be additional "susceptible" sub-groups to consider. Health research in recent years has raised CNS and brain impacts as a health outcome of air pollution concern, as well as pregnant mothers and early-life exposure of the fetus and infant as especially susceptible issues, for example. Additionally, 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, and other socio-demographic factors should also be addressed; (in other words, there should not be a presumption that the three sub-groups previously identified are the only needed foci).

P2-4, mid-page (2nd dash heading, 4th bullet) - I welcome and strongly endorse the consideration of "co-occurring" 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)? The Clean Air Act directs that reviews be focused on the pollutant or indicator of concern, but science has shown quite clearly that multi-pollutant exposures do lead to meaningful (and sometimes more substantive) decrements in health than consideration of single-pollutant exposures. Nitrogen oxides is potentially one of those pollutants for which multi-pollutant exposures may be problematic.

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 population, exposure, comparison, outcome, study 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?

In a similar manner, the satellite imagery and possibly the deployment and use of low-cost sensor networks operated by the public and uploaded to viewable websites have steadily increased since the last review. Are there any additional exposure databases that could be exploited to gain improved understanding of NO2 trends?

P3-25, Table 3-8: this summary of causal determinations from the 2016 ISA reminds us that both life stage AND health outcomes are of critical importance. What (if any) information has become available in the past five years that would inform us about neurologic or respiratory declines in older adults? What about neurologic/cognitive development effects in pregnancy and early childhood?

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) Occupations potentially associated with elevated NO2 exposures (such as mining operations, railyard, or cargo-goods movement related exposures) might be informative for health outcomes of possible importance, even if the exposures may not be deemed "ambient". 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. Judith Chow

Volumes I and II of the IRP present a good overview of the history and future perspectives for the upcoming NAAQS review for oxides of nitrogen. More discussion is needed to clarify: 1) nitrogen dioxide (NO_2) vs oxides of nitrogen; 2) near-road monitoring data; and 3) the scope of the Integrated Science Assessment (ISA).

NO₂ vs Oxides of Nitrogen

Although the IRP calls for primary NAAQS for oxides of nitrogen, both Volumes I and II of the IRP focus on the NO_2 indicator. As a large number of oxides of nitrogen species are emitted into the atmosphere, sources and concentration levels of other reactive and nonreactive nitrogen compounds that participate in atmospheric transformations and transport also contribute to the adverse health and ecosystem effects. Studies related to nitrate radicals (NO_3), nitric acid (HNO_3), dinitrogen pentoxide, peroxyacetyl nitrate (PAN), and others should be considered.

IRP Volume I (footnote 15, p 3-5) states that NO_2 is most relevant to the evaluation of health evidence without sufficient rationale. As section 108(c) of the CAA (quoted in footnote 16, p. 3-5) specifies that oxides of nitrogen "…shall include a discussion of nitric and nitrous acids, nitrites, nitrates, nitrosamines, and other carcinogenic and potentially carcinogenic derivatives of oxides of nitrogen." This statement is reiterated in IRP Volume II (footnote 14, p. 3-3). A comprehensive diagram that illustrates nitrogen fates and transport in the atmosphere would clarify the relationships among the different reactive and non-reactive compounds.

Section 3.4.1 (Source to Concentration, p 3-20) proposes to apply new modeling methods to better understand the spatial and temporal patterns of NO_2, and more broadly NO_y (total reactive nitrogen compounds, including NO, NO_2, HNO_3, PAN, and others). NO_y measurements are available from EPA's NCore, PAMS, and CASTNET networks. Data analysis should include NO, NO_2, and NO_y concentrations available from these networks to complement and verify the modeling results.

Near-Road Monitoring Data

Although multi-pollutant air quality management is acknowledged, Section 3.4.1 (p. 3-20) in Volume II (Source to Concentrations) emphasizes NO_2 over the other nitrogen species. The IRP Volume I states that there were 75 near-road monitors operating in 2022-2023. However, only 64 sites are summarized in Table A-3. The U.S. EPA near-road website (https://www.epa.gov/amtic/near-road-monitoring) shows that NO_2 and CO are most commonly measured at these sites, with some sites also quantifying $PM_{2.5}$ mass, black carbon (BC), ultrafine particles, ozone, and meteorology. Distances of measurements from the roadside vary from 7 to 85 meters. Detailed analyses of the spatial variability and near-road increments are needed. Data from the near-road sites (https://www.epa.gov/outdoor-air-quality-data) could be useful for evaluating exposures to nearby under-represented communities. Additional analysis of the data acquired to date, with consideration of the site descriptions, roadway conditions, and distance from the roadside would be a useful contribution.

Scope of the ISA Review (Section 3.2, Volume II)

The cutoff date for the cited literature was March 2014 for the 2016 oxides of nitrogen ISA. The upcoming ISA has a cutoff date of June 2023. This cutoff date may ignore new evidence that is published between June 2023 and the issuance of the final ISA. The IRP states that review articles that summarize and interpret existing studies rather than including new data or analyses will be excluded. Several of these articles can be relevant when they synthesize the available information, evaluate the pros and cons of past studies, and identify lessons learned for future studies. These reviews should not be a priori excluded from consideration.

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

The draft document specified PECOS exposure criteria for epidemiologic studies for annual averages in comparison to the 98th percentile of annual average NO_2 concentrations measured at ambient air monitors in the U.S. As mentioned during the meeting, this was a method to flag studies that used ambient concentrations typical of US NO_2 data. That reasoning should be more clearly stated and valid studies should not be eliminated based on concentration.

Are there any other concentration or averaging period restrictions on studies looking at the annual or shorter-term exposures? NO_2 is seasonal in the Northeast and at urban and near-road sites, higher concentrations typically occur for a few hours on weekday mornings.

Are motor-vehicle tunnel or toll booth studies being examined to look for areas with higher concentrations?

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 median 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?

There are now instruments available using technologies such as CRDS that are more precise and accurate than current FRMs. Will there be an analysis of the adequacy of the current FRM and FEMs deployed for the various users of the National NO and NO_2 dataset?

The draft document does not indicate that the NO_2 monitoring network design will be evaluated. The adequacy of the network should always be evaluated so that it can be adjusted to meet future needs.

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. Mark Frampton

<u>General Comments</u>

Volume 2 of the Integrated Review Plan (IRP) for the Primary National Ambient Air Quality Standards for Oxides of Nitrogen represents a thorough, comprehensive, clearly presented document laying the groundwork and approaches for the scientific review of the health effects of exposure to nitrogen oxides, and preparation of the associated Integrated Science Assessment (ISA).

The document builds on previous nitrogen oxide ISAs, with appropriate focus on evidence that has become available after the cutoff date for the prior review. The approach takes into consideration previous CASAC recommendations, and findings and recommendations from the ad hoc committee of the National Academy of Sciences, Engineering, and Medicine (NASEM).

The overall organization of this IRP is well considered and logical, following approaches taken in prior IRPs, building on their strengths. The organization of the ISA will use the term "chapters" to describe the topical divisions, a change from "appendices" used in recent ISAs. This is appropriate and appreciated, allowing the major sections of the ISA to be distinguished from supplementary material.

The document is supported by an Appendix that reviews in more detail the methods and approaches used in preparing ISAs in general.

<u>Specific Comments</u>

2.1 REVIEW OF THE PRIMARY NO$_2$ STANDARDS

The listing of questions to be addressed in the ISA should be consolidated and reorganized. For example on page 2-4, the sub-paragraphs under the second bullet ("To what extent is key scientific evidence available to improve or alter our understanding of the health effects associated with various time periods…") do not deal with time periods. They are separate questions that should be advanced one level. The first two sub-bullets under this heading deal with concentrations and say essentially the same thing; they could be merged.

It is unclear why the last sub-bullet in this section ("To what extent is new information available to improve the characterization of the severity and/or potential adversity of NO$_2$-induced respiratory effects reported in controlled human exposure studies? To what extent does such information inform an understanding of effects in at-risk populations?") focuses specifically on respiratory effects. It seems the question is applicable to all potential health effects that might be addressed in human exposure studies. And the reference to at-risk populations duplicates the section that follows.

3.2 Scope of the ISA.

This ISA will be appropriately limited to studies with new information since the previous ISA, and that have undergone peer review. Studies considered will include a full range of potential health effects, including populations and lifestages at increased risk, with integration with findings from the 2016 ISA.

Of note, according to Table 3-1 listing the PECOS criteria for epidemiology, studies considered will not be limited geographically.

Footnote 16 on page 3-5 indicates a concentration cutoff of "overall average NO_2 exposures … at or below 22 ppb," noting this is the 98th percentile of annual average NO_2 concentrations in the U.S., 2008-2022. It would seem more relevant to use a cutoff equivalent to the current U.S. annual NAAQS of 53 ppb. And presumably the cutoff should be higher for short-term exposure studies. According to https://www.epa.gov/system/files/documents/2023-06/NO2_2022.pdf, the 98th percentile MDA1 for 2020-2022 ranged from 38.7 to 49.0. The document should either indicate that this cutoff applies only to long-term exposure studies, or provide a separate, higher cutoff for short-term exposure studies. And it is suggested that these exposure limits be included in the relevant sections of the PECOS tables, rather than relegated to a footnote.

Page 3-23, second bullet is unclear. "Relative contributions to local NO_2..." seems to ask about sources for local exposures, but then veers into dietary sources and endogenous formation. Perhaps there is a need to define "local" here; perhaps it is meant to indicate personal NO_x burden. If that is the case, one needs to consider that endogenous NO production is often physiologically beneficial. This bullet needs clarification.

Appendix, A.5.2. Controlled Human Exposure

This section of the Appendix provides an excellent general discussion of the design issues, strengths, and weaknesses of controlled human exposure studies (CHE). However, the discussion would benefit from additional cautions about interpretation of studies with "negative" findings.

1. For CHE studies, the observed effects are applicable and causal only for people with characteristics similar to the exposed participants, and under the specific circumstances of the study. These participants are generally healthier and more fit than the general population, and do not include children, the very elderly, or those who are frail or severely obese. Minority groups are often under-represented. Potential participants who have difficulty following instructions or performing the maneuvers needed for outcome testing will generally be excluded. Yet these are the individuals that may be at increased risk. Participants in CHE studies may be less likely than the general population to demonstrate pollutant effects on the biomarkers being measured.
2. Exposure to ambient pollutants in the minutes or hours prior to the experimental exposure, for example in traffic on the way to the laboratory, may affect the outcome measures of the study, especially if the study concentrations are relatively low. This added variability may bias the result towards the null.
3. There is also a need to carefully consider the degree to which the outcomes being measured capture the adverse health effects of importance. For example, a decrease in lung function, as measured by FEV1, is considered adverse, because reductions in this indicator are associated with poorer outcomes in people with respiratory disease. A transient increase in blood pressure may be appropriately considered adverse because of the known adverse cardiovascular effects of hypertension. However, exercise transiently increases blood pressure, but is beneficial to cardiovascular health overall. Adverse respiratory effects can occur without a reduction in FEV1, especially if the exposure predominantly affects the small airways. And pollutant exposures may adversely affect cardiovascular health without acutely affecting blood pressure. CHE studies, by design, should not elicit clinically adverse health effects, such as asthma exacerbations,

respiratory infections, or myocardial infarctions. Thus, while it is true that CHE studies assist in determining causality because of their controlled design, proving causality for a change in a biomarker may not be the equivalent of proving causality for a disease state.

4. In addition, delayed effects can be missed; most CHE studies do not assess effects beyond 24 hours after exposure.

For these reasons, caution is needed in using CHE studies to determine absence of health effects, no-effect levels, or threshold concentrations.

<u>Minor Comments</u>

The last paragraph on page 3-2 contains duplicate sentences.

The first paragraph on page A-1 contains a repeated series of sentences

Dr. Christina H. Fuller

Consistency in at-risk terminology

The appendix includes a good discussion on the variability in terminology within the scientific community to describe populations that may be at increased risk of adverse effects of exposure due to specific factors (pages A-51 – A-52). On page A-51 is a description of the history behind the EPA's decision to utilize the term "at-risk" for these populations and lifestages. Among populations/lifestages described as at-risk there are specific groups of factors and there are some inconsistencies in the document that need to be corrected. I list those that I have noticed here. In footnote 11 on page 2-3 at-risk groups are divided into susceptible and vulnerable; while on page 3-29 at-risk groups are described as intrinsic, acquired, extrinsic and factors affecting dose or exposure. On page A-51, at-risk groups are listed as intrinsic, extrinsic and exposure-related. I suggest dividing factors into three groups and ensuring that the language is consistent throughout (i.e. intrinsic, extrinsic and exposure-related). Also necessary, evaluate interactions between factors such as intrinsic (e.g. age) and exposure-related (e.g. school next to a busy highway). The importance of assessing these types of interactions is noted on page A-51.

Include new data from already identified at-risk groups

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 intrinsic, extrinsic or exposure-related factors. Although this may be the plan of EPA staff it is important to specify in the figure.

Integrate at-risk populations into presentation of health findings

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 at-risk groups into the main health assessment portions of the document. Since the purpose of the standards is to ensure protection of health for the general population **and** sensitive groups (i.e., at-risk) presenting both in the same section allows for more straightforward assessment. The data is easier to understand this way rather than to separate the at-risk discussion into a section detached from the other health outcomes.

Delineate at-risk factors in tables and throughout the document

Table 3-9. The categories of factors within at-risk (i.e., intrinsic, extrinsic, and exposure-related) should be delineated in this table and others where relevant. These factors are not all conceptually the same nor are they evaluated in the same way, so delineation here increases clarity. For example, there is suggestive evidence for female sex (an intrinsic factor) as well as reduced antioxidant intake and low SES (extrinsic factors).

PECOS criteria and at-risk groups

Application of PECOS criteria is positive for conducting structured identification of studies for inclusion in the ISA. It is also important to consider if any of the PECOS criteria may disproportionately disqualify high-quality studies that address exposures and outcomes to at-risk populations.

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. Unless there are frequent changes, it might 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 unnecessary. 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 NO_2 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. Petros Koutrakis

The plan presented in Volume 2 is comprehensive and I do not have any comments. This approach has been established for previous reviews so should be fine.

My only comment has to do with the relationship between nitrogen oxides and ozone levels. Ozone poses a serious health risk as well, and its levels have not decreased much despite the emission reductions of its precursors. Of course, there are other important factors such as climate change or increase in biogenic emissions. However, a thorough literature review about linkages between nitrogen oxides and ozone levels should be included.

Dr. Julian Marshall

- In multiple places[1], the report lists factors that may lead to differential exposures, including proximity to sources, activities patterns, and socioeconomic status. On page 3-29, attributes are listed more explicitly (Table 3-9). The term "socioeconomic" generally excludes race/ethnicity. For exposure to NO2, there is evidence that race matters independently of income.
    - Does EPA plan to include race/ethnicity in the list of factors and, more broadly, in the document in general? Exposure disparities are not just socioeconomics; race/ethnicity matters too. I did not see race/ethnicity mentioned in the report.
    - In one place (p. 3-30), the report uses the term "sociodemographic" instead of "socioeconomic". Perhaps make this a general switch? "Sociodemographic" is more general and (in my mind) *includes* the latter term; if that's true, would "sociodemographic" be preferable to "socioeconomic"?
- The report talks in multiple places (e.g., p. 3-21) about new methods and the resulting newly available information; examples given include low-cost sensors, community monitoring, satellite data, LUR, and data fusion.
    - What does that new information, and the new methods now available, tell us about where to locate ambient monitors?
    - Do the new methods shed light on the degree to which the current NAAQS protect more-susceptible / more-vulnerable populations?
    - Mobile-monitoring should be added to the list of new methods.
    - Also, perhaps add "hybrid methods" or "combined method" to the list, to reflect approaches that combine multiple approaches. (While the term "data fusion" may overlap with "hybrid" / "combined" methods, often "data fusion" research sometimes is more about the fusion techniques themselves.)
- I could not access the database of articles (page 3-13 / Figure 3-1).

---

[1] Examples:

[page 2-4] "To what extent is new information available to improve our understanding of the NO2 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?"

[Figure 2-1, page 2-7] "Newly identified at-risk populations"

[Table 3-5, page 3-10] "exposure determinants (i.e., factors which may lead to differential exposures, such as proximity to sources, activity patterns, and socioeconomic status),"

[3-22] "What implications do potential differences in exposure measurement error have on inferences about relationships with health effects observed in general population studies versus those conducted in specific lifestages and groups within the population (e.g., people with underlying health condition)?"

References that May be of Interest

Disparities in exposure to NO2

J Liu, LP Clark, MJ Bechle, A Hajat, SY Kim, AL Robinson, L Sheppard, AA Szpiro, JD Marshall. Disparities in air pollution exposure in the United States by race/ethnicity and income, 1990–2010. Environmental Health Perspectives. 2021, 129(12), DOI: 10.1289/EHP8584.

MJ Bechle, DB Millet, JD Marshall. Ambient NO2 air pollution and public schools in the United States: relationships with urbanicity, race–ethnicity, and income. Environmental Science & Technology Letters. 2023, 10, 844-850, DOI: 10.1021/acs.estlett.3c00507.

LP Clark, DB Millet, JD Marshall. National patterns in environmental injustice and inequality: outdoor NO2 air pollution in the United States. PLOS ONE. 2014, 9(4), e94431.

LP Clark, MH Harris, JS Apte, JD Marshall. National and intraurban air pollution exposure disparity estimates in the United States: impact of data-aggregation spatial scale. Environmental Science & Technology Letters. 2022, 9(9), 786-791, DOI: 10.1021/acs.estlett.2c00403.

LP Clark, DB Millet, JD Marshall. Changes in transportation-related air pollution exposures by race-ethnicity and socioeconomic status: outdoor nitrogen dioxide in the United States in 2000 and 2010. Environmental Health Perspectives.2017, 125(9):097012, DOI: 10.1289/EHP959.

Dressel, I. M., Demetillo, M. G. D., Judd, L. M., Janz, S. J., Fields, K. P., Sun, K., Fiore, A. M., McDonald, B. C., and Pusede, S. E.: Daily satellite observations of nitrogen dioxide air pollution inequality in New York City, New York and Newark, New Jersey: Evaluation and application, Environ. Sci. Technol., 56, 22, 15298–15311, 2022

Demetillo, M. A. G., Harkins, C., McDonald, B. C., Chodrow, P. S., Sun, K., and Pusede, S. E.: Space-based observational constraints on NO2 air pollution inequality from diesel traffic in major U.S. cities, Geophys. Res. Lett., 48, e2021GL094333, 2021

Demetillo, M. A. G., Navarro, A., Knowles, K. K., Fields, K. P., Geddes, J. A., Nowlan, C. R., Sun, K., Judd, L. M., Al-Saadi, J., Diskin, G. S., McDonald, B. C., and Pusede, S. E.: Observing nitrogen dioxide air pollution inequality using high-resolution remote sensing measurements in Houston, Texas, Environ. Sci. Technol., 54, 16, 9882–9895, 2020

Disparities in air pollutants across racial, ethnic, and poverty groups at US public schools
MJ Cheeseman, B Ford, SC Anenberg, MJ Cooper, et al. GeoHealth 6 (12), e2022GH000672

Increasing racial and ethnic disparities in ambient air pollution-Attributable morbidity and mortality in the United States. GH Kerr, A van Donkelaar, RV Martin, M Brauer, K Bukart, S Wozniak, et al. Environmental health perspectives 132 (3), 037002

<u>Modeling concentrations and spatial patterns in NO2</u>

M Qi, K Dixit, JD Marshall, W Zhang, S Hankey. National land use regression model for NO2 using street view imagery and satellite observations. Environmental Science & Technology. 2022, 56(18), 13499-13509, DOI: 10.1021/acs.est.2c03581.

Y Wang, MJ Bechle, SY Kim, P Adams, SN Pandis, CA Pope III, AL Robinson, L Sheppard, AA Szpiro, JD Marshall. Spatial decomposition analysis of NO2 and PM2.5 air pollution in the United States. Atmospheric Environment. 2020, 241, 117470, DOI: 10.1016/j.atmosenv.2020.117470.

MT Young, MJ Bechle, PD Sampson, AA Szpiro, JD Marshall, L Sheppard, JD Kaufman. Satellite-based NO2 and model validation in a national prediction model based on universal Kriging and land-use regression. Environmental Science & Technology. 2016, 50(7), 3686–3694.

K de Hoogh, J Gulliver, A van Donkelaar, RV Martin, JD Marshall, MJ Bechle, G Cesaroni, MC Pradas, A Dedele, M Eeftens, B Forsberg, C Galassi, J Heinrich, B Hoffmann, B Jacquemin, K Katsouyanni, M Korek, N Kunzli, SJ Lindley, J Lepeule, F Meleux, A de Nazelle, M Nieuwenhuijsen, W Nystad, O Raaschou-Nielsen, A Peters, VH Peuch, L Rouil, O Udvardy, R Slama, M Stempfelet, EG Stephanou, MY Tsai, T Yli-Tuomi, G Weinmayr, B Brunekreef, D Vienneau, G Hoek. Development of West-European Personal PM2.5 and NO2 land use regression models incorporating satellite-derived and chemical transport modelling data. Environmental Research. 2016, 151, 1-10.

MJ Bechle, DB Millet, JD Marshall. National spatiotemporal exposure surface for NO2: monthly scaling of a satellite-derived land-use regression, 2000 - 2010. Environmental Science & Technology. 2015, 49(20), 12297–12305.

D Vienneau, K de Hoogh, MJ Bechle, R Beelen, A van Donkelaar, RV Martin, DB Millet, G Hoek, JD Marshall. Western European land use regression incorporating satellite- and ground-based measurements of NO2 and PM10. Environmental Science & Technology. 2013, 47(23), 13555–13564.

MJ Bechle, DB Millet, JD Marshall. Remote sensing of exposure to NO2: satellite versus ground-based measurement in a large urban area. Atmospheric Environment. 2013, 69, 345–353.

EV Novotny, MJ Bechle, DB Millet, JD Marshall. National satellite-based land-use regression: NO2 in the United States. Environmental Science & Technology. 2011, 45(10), 4407–4414.
(This study is, to my understanding, the first national LUR model for the US, and so is the first national LUR in the US for NO2.)

LP Clark, MH Harris, JS Apte, JD Marshall. National and intraurban air pollution exposure disparity estimates in the United States: impact of data-aggregation spatial scale. Environmental Science & Technology Letters. 2022, 9(9), 786-791, DOI: 10.1021/acs.estlett.2c00403.
(this citation is included above too, under 'disparities in exposure'; it applies to both categories.)

MJ Bechle, ML Bell, DL Goldberg, S Hankey, T Lu, AA Presto, AL. Robinson, J Schwartz, L Shi, Y Zhang, JD Marshall. Intercomparison of six national empirical models for PM2.5 air pollution in the contiguous US. Findings. 2023, November, DOI: 10.32866/001c.89423.
(This paper is about PM2.5, not NO2, but I think the findings may still potentially have relevance. They intercompare six very different national empirical models for PM2.5, and find that the resulting

estimates are relatively highly correlated. We don't know with certainty what a similar comparison, but for NO2, would look like; this result suggests it's possible that national NO2 models are similar to each other.)

T Lu, JD Marshall, W Zhang, P Hystad, SY Kim, MJ Bechle, M Demuzere, S Hankey. National empirical models of air pollution using microscale measures of the urban environment. Environmental Science & Technology. 2021, 55(22), 15519-15530, DOI: 10.1021/acs.est.1c04047.

SY Kim, MJ Bechle, S Hankey, L Sheppard, AA Szpiro, JD Marshall. Concentration of criteria pollutants in the contiguous U.S., 1979 - 2015: Role of prediction model parsimony in integrated empirical geographic regression. PLOS One. 2020, 15(2): 0228535, DOI: 10.1371/journal.pone.0228535.

LD Knibbs, CP Coorey, MJ Bechle, JD Marshall, MG Hewson, B Jalaludin, GG Morgan, AG Barnett. Long-term nitrogen dioxide exposure assessment using back-extrapolation of satellite-based land-use regression models for Australia. Environmental Research. 2018, 163:16-25, DOI: 10.1016/j.envres.2018.01.046.

MN Blanco, A Gassett, T Gould, A Doubleday, D Slager, E Austin, E Seto, T Larson, JD Marshall, L Sheppard. Characterization of annual average traffic-related air pollution levels (particle number, black carbon, nitrogen dioxide, PM2.5, carbon dioxide) in the greater Seattle area from a year-long mobile monitoring campaign. Environmental Science & Technology. 2022, 56(16), 11460-11472, DOI: 10.1021/acs.est.2c01077.

A Larkin, JA Geddes, RV Martin, Q Xiao, Y Liu, JD Marshall, M Brauer, P Hystad. A global land use regression model for nitrogen dioxide air pollution. Environmental Science & Technology. 2017, 51(12), 6957–6964, DOI: 10.1021/acs.est.7b01148.

L Knibbs, C Coorey, MJ Bechle, C Cowie, M Dirgawati, J Heyworth, G Marks, JD Marshall, L Morawska, G Pereira, M Hewson. Independent validation of national satellite-based land-use regression models for nitrogen dioxide using passive samplers. Environmental Science & Technology. 2016, 50(22), 12331–12338.

Health effects of NO2

PF Coogan, LF White, J Yu, RT Burnett, JD Marshall, E Seto, RD Brook, JR Palmer, L Rosenberg, M Jerrett. Long term exposure to NO2 and diabetes incidence in the black women's health study. Environmental Research. 2016, 148, 360-366.

R Alotaibi, MJ Bechle, JD Marshall, T Ramani, J Zietsman, M Nieuwenhuijsen, H Khreis. Traffic related air pollution and the burden of childhood asthma in the contiguous United States in 2000 and 2010. Environment International. 2019, 127, 858-867, DOI: 10.1016/j.envint.2019.03.041.

Sensitivity of estimated NO2-attributable pediatric asthma incidence to grid resolution and urbanicity. A Mohegh, D Goldberg, P Achakulwisut, SC Anenberg. Environmental Research Letters 16 (1), 014019

Changes in NO2 during COVID

J Xiang, E Austin, T Gould, T Larson, J Shirai, Y Liu, JD Marshall, E Seto. Impacts of the COVID-19 responses on traffic-related air pollution in a Northwestern US city. Science of The Total Environment. 2020, 747, 141325, DOI: 10.1016/j.scitotenv.2020.141325.

B Bekbulat, JS Apte, DB Millet, AL Robinson, KC Wells, JD Marshall. Changes in criteria air pollution levels in the US following societal Covid-19 response: evidence from regulatory monitors before, during, and after stay-at-home orders. Science of the Total Environment. 2021, 769, 144693, DOI: 10.1016/j.scitotenv.2020.144693.

COVID-19 pandemic reveals persistent disparities in nitrogen dioxide pollution
GH Kerr, DL Goldberg, SC Anenberg. Proceedings of the National Academy of Sciences 118 (30), e2022409118

Disentangling the Impact of the COVID-19 Lockdowns on Urban NO2 From Natural Variability
DL Goldberg, SC Anenberg, D Griffin, CA McLinden, Z Lu, DG Streets
Geophysical Research Letters 47 (17), e2020GL089269

Dr. Michele 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. Jennifer Peel

Comments for Volume 2

1. Page 2-4: Given the recent CASAC discussion for PM and Ozone, I would like to see more clarification of how the EPA is approaching these questions. I encourage the EPA to revisit recent recommendations from the CASAC, overall and specifically with regard to the questions posed below (and in #2).

   "To what extent is key scientific evidence available to improve or alter our understanding of the health effects associated with various time periods of exposures, including short-term (e.g., 1-hour) and long term exposures (e.g., more than one month to years) to oxides of nitrogen?
     o At what pollutant concentrations do these health effects occur?
     o To what extent is new information available to improve our understanding of the range of ambient air concentrations within which oxides of nitrogen contribute to health effects?
     o Is there evidence of effects at oxides of nitrogen exposure concentrations lower than those at which effects were previously observed or in areas that would likely have met the current primary NO_2 standards?"

2. And similarly, from Pages 2-4 and 2-5:

   "To what extent does risk or exposure information suggest that exposures of concern are likely to occur with recent ambient air NO_2 concentrations in the U.S. or with concentrations that just meet the current primary NO_2 standards?
     o Are the estimated exposures/risks considered in this review of sufficient magnitude such that the health effects might reasonably be judged to be important from a public health perspective?
     o What new information is available to improve our understanding of exposure measurement error and the role of exposure in epidemiologic inference, particularly for interpreting long-term exposure studies?
     o What are the important uncertainties associated with any exposure/risk estimates?"

3. Table 3-1 includes a limited set of outcomes in the PECOS statement; the text on the preceding page (Page 3-4) includes language that indicates that all health outcomes will be evaluated, not just the ones in Table 3-1. Please clarify what outcomes will be included. I encourage EPA to include a broader range of health outcomes for consideration, including but not limited to, neurologic, reproductive, renal. Thes same comment applies to all of the PECOS statements in Volume 2. It appears that the search conducted (Page 3-11) was sufficiently comprehensive to include this broader range of outcomes, but I do not see the list of search terms. Section 3.3.2 makes it sounds like the screening process of titles and abstracts did indeed use the PECOS criteria (including the limited set of Outcomes). I do not think this is the case because Table 3-6, which is stated to include the results of the screening process, includes a range of outcomes. Please clarify.

4. Section 3.3.3 uses a process similar to that of past reviews and that has been reviewed and supported by the NASEM review. It would be helpful to clarify footnote 22 on page 3-15:

"[22]Specifically, models used to estimate oxides of nitrogen exposures in epidemiologic studies (e.g., land use regression models or ensemble machine learning models) should be validated for the location(s) and population(s) under investigation. The lack of such validation was an important uncertainty in some epidemiologic studies evaluated in the last review of the primary NO2 NAAQS (e.g., 83 FR 17268, April 18, 2018)."

Validation can entail various approaches that have a range of interpretations in terms of the extent of validity of the model estimates. I'm not sure that any exposure model is 100% "validated" in the sense that there is no uncertainty, and the validation / evaluation of the model is only relative to what it is being compared.

5. Similarly, footnotes 23 and 24 could use some additional discussion / clarity.

   "[23]Studies described as "natural experiments" conducted during COVID lockdown(s) should explicitly consider potential confounders common during lockdown periods, such as changes in employment status, activity patterns (e.g., time spent outdoors versus indoors, driving, working, exercising), stress levels, access to health care, and/or mask wearing.

   [24]Specifically, epidemiologic studies that examine populations outside North America should be multicity and/or multi-country and they should address policy-relevant topics (e.g., studies that use alternative methods for confounder control (causal inference methods, quasi-experimental studies), copollutant confounding, effect measure modification for potential at-risk factors (race/ethnicity, age, SES indicators, etc.), exposure-/concentration-response relationships). Studies with these characteristics are most likely to be influential in ISA causality determinations and other conclusions. Studies that examine populations in North America will not be excluded from the ISA based on these study characteristics alone as such studies may be useful for evaluating potential policy options in subsequent steps of the NAAQS review."

   While I understand the reason and need for criteria to screen out some of the identified relevant articles, the process of screening out articles that will likely not influence the ISA's conclusions, and therefore not considering them further in the ISA conclusions and at the next steps in the process, may exclude some studies that could have been helpful for the process – e.g., for health outcomes for which there is less evidence from North America, for consideration of at-risk groups, or for evaluation of evidence within a certain range of pollutant concentrations (especially at low concentrations).

6. The questions posed on pages 3-27 and 3-28, for short-term and long-term exposure, seem to be focused on a limited number of health outcomes. Why is this the case?

7. Recent ISAs (e.g., PM and ozone) have updated language with regarding to potentially at-risk populations; I encourage EPA to consider using the updated language and approach, particularly when discussing considering individual characteristics, location and place, gender identity, race, economic indicators, etc.

8. Please consider the language in the causal determinations that seems to be different for Causal and Likely to Be Causal in terms of certainty of ruling out other explanations.

Dr. Richard Peltier

Minor Comments

1. 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.
2. 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..."

Other Comments

1. Chapter 3 (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[2] and could pose an important source of NOx exposure where few primary emissions were expected.
2. 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. Like ozone, NO2 is a relatively pure chemical and an indicator for a class of photochemical oxidants, and it likely have similar causal pathways across different populations.
3. It would 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?
4. 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 where high prevalence use of indoor gas cookstoves leads to increased indoor NOx concentrations[3] and exposures, does this lead to a population that is increasingly susceptible or sensitive to adverse health outcomes under more modest ambient NOx conditions? I would point EPA to these two references for consideration.

---

[2] Aruffo et al, EAS&T, 2022
[3] E.g. Zhao et al, Indoor Air 2020

Krasner, Andee, M.P.H., et al. "Cooking with Gas, Household Air Pollution, and Asthma: Little Recognized Risk for Children." *Journal of environmental health* 83.8 (2021): 14-8.

Weiwei Lin, Bert Brunekreef, Ulrike Gehring, Meta-analysis of the effects of indoor nitrogen dioxide and gas cooking on asthma and wheeze in children, *International Journal of Epidemiology*, Volume 42, Issue 6, December 2013, Pages 1724–1737

5. On a related note, EPA uses the term 'cookstove' (Appendix A) to describe major indoor sources of NO2 exposure. In reviewing of the literature, EPA will likely find that the vast majority of retrieved manuscripts are likely to involve primitive cookstoves (e.g. three-stone stoves, improved biomass stoves, or a range of propane intervention approaches – most of which are generally not found in the United States). It is important to better define the term 'cookstove' to include modern piped gas stoves and assess the potential confounding impact of this source; EPA should remain mindful of this in their literature reviews.

6. 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. To be clear, I recommend continuing to use 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. The cited references were simply validations of these tools using assessments that are not related to air quality, and it is important to demonstrate that these are robust and lacking in bias

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.

CHAPTER 2

1. PAGE 2-2: "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 2-3: "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 2-4: "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 it's just a mean or median or a range, since those are most often reported, but please make it clear.

4. PAGE 2-4: "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?

CHAPTER 3

5. PAGE 3-27: "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 should be added here to clarify this approach.

6. PAGE 3-26: "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?

7. PAGE 3-27 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?

8. PAGE 3-28: "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? A trimester average is considered long term in the document? For studies of pregnancy health effects, this is an important consideration given the expanded use of distributed lag models in such studies. In the analysis, a prior time period (e.g., gestational month 1, gestational weeks 14-20) may not be defined in the statistical analysis plan, but just a specific set of weeks may be identified using this analytic procedure. This period may be very short (e.g., gestational week 9) or much longer (e.g., gestational weeks 17-35). As we discussed,

a statement within each outcome group should be made as to what defines a 'short-term' and 'long-term' exposure. 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. Therefore, deciding on a process by which you will make these decisions in this ISA should be made, so that future ISA's that will likely have a much larger number of pregnancy epidemiology studies done this way can.

APPENDIX

1. PAGE A-25: "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?

2. PAGE A-26: "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.

Dr. Elizabeth A. (Lianne) Sheppard

<u>IRP Volume 2</u>

General comments

1. Overall, the IRP Volume 2 is a helpful document that provides a good foundation for the ISA review. I commend the EPA staff for the strong first draft and the preliminary work already completed. The IRP will be further improved by some revisions 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. This was clearly presented during the EPA briefing and a basic summary of these plans should be articulated in the IRP
3. As some other panelists have mentioned, 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 should become standard practice to incorporate into Chapter 2 a discussion of the previous CASAC advice provided in previous review(s), EPA's decisions in response to that advice, and the EPA's plans to address that previous advice, if any, in the current review. (Also consider subheadings for each of those 3 topics as part of the standard outline for Chapter 2.) Of particular note, if there was any discrepancy between the CASAC recommendations and the EPA final decision in the past review, or in the case that 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. This will support continuity across reviews and explicitly recognize that the experts on CASAC change over time such that this summary will support their work.
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? I think at least an anticipated schedule should be outlined, even if a court decision may necessitate changes.

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 the EPA should 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 (and/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? As we learned from the EPA at the NOx IRP review meeting, the 2015 ISA Preamble is out of date and there are plans for it to be updated after multiple revisions and CASAC reviews. Appendix A is the current draft. This is a solid plan and it should be articulated in the document. 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? The "e.g.," refers to various health outcomes, but not how they are quantified.
- 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? I think the workshop and its purpose should be mentioned and put into context with the entire process.

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? The discuss and clarification of this point during the NOx IRP review meeting would be valuable to incorporate into the document.

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

P 3-17 last sentence: As noted above, I think the EPA should articulate their plan for updating the Preamble in this document.

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. In the NOx IRP review meeting, the EPA staff stated that if the CASAC recommends a second draft, then this can happen. This should be stated explicitly in the IRP.

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, e.g. in a preamble to the 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 necessarily designed to address different questions. For instance, there are no panel studies of a subject such as out of hospital cardiac arrest, which could be considered in a case-crossover study. I think the blanket 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-20 Section A.5.1:

- P A-21: I concur with Dr. Zigler's lack of agreement with the following statement: "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." His comments provide the rationale.
- P A-21: I don't agree with the statement that case-crossover (and panel) studies should be prioritized over time series studies. The seminal paper introducing the time-stratified case-crossover design, Lumley & Levy (2000 Envirometrics), indicates that case-crossover studies use the same estimating functions as time series studies; the distinction is that case-crossover studies use matching rather than modeling to address temporal confounding. Panel studies, while useful, are typically very small and have many of the same study population and endpoint-related limitations as CHEs.
- P A-23 to A-24: These statements about spatio-temporal variability of exposure are too broad. They are relevant to time series and other short-term exposure studies that rely on exposure time series, as do the references cited, which focus on the time series design. Considerations about spatio-temporal variability of exposure are different for cohort studies and are not addressed in this text. This text should also clarify what validation aspects would give a study greater versus less weight.
- P A-25: For application to epidemiologic inference:
  - I strongly favor exposure models that are based on exposure measurements rather than those based exclusively on models such as chemical transport models which are based on understanding (i.e., conceptual models) of the physical and chemical understanding of air pollution that are not necessarily grounded in or validated by models.
  - Population stochastic models are appropriately applied in risk assessment analyses, but NOT for epidemiologic inference.

P A-30 Section A.5.2: I wish to echo and strengthen Dr. Frampton's comments about the limitations of CHEs. While CHEs are very useful in determining the presence of an effect is causal, it is inappropriate to use them to conclude that an effect is absent when these studies fail to detect an effect. Absence of evidence to show an effect is not evidence that such an effect is absent.

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. I think this workshop (and the one prior to development of the ISA, if that one is distinct) should be added to the milestones table.

P A-55-A-56: There is no mention of a possible second draft ISA. Please see my previous comment on this point and make sure that the document does not close the door to second drafts.

IRP Volume 1

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 of folding the REA analyses into the PA 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. This should be rectified.

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. This should be made explicit in the IRP, particularly given that EPA staff have said in public meetings, including the recent NOx IRP meeting, that they will produce second drafts when they are warranted. This conflicts with the IRP text.

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 at 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. 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 de- signs. Along these same lines, there is an increasing trend in clinical

A-42

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 pack- aging 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 improve 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 road- way 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 Flaxman. 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. Hernán 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.