UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON D.C. 20460

OFFICE OF THE ADMINISTRATOR
SCIENCE ADVISORY BOARD

September 9, 2015

EPA-CASAC-15-001

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

Subject: CASAC Review of the EPA's *Integrated Science Assessment for Oxides of Nitrogen – Health Criteria (Second External Review Draft – January 2015)*

Dear Administrator McCarthy:

The Clean Air Scientific Advisory Committee (CASAC) Oxides of Nitrogen Primary National Ambient Air Quality Standards (NAAQS) Review Panel met on June 2, 2015, and August 13, 2015, to peer review the EPA's *Integrated Science Assessment for Oxides of Nitrogen – Health Criteria (Second External Review Draft – January 2015)*, hereafter referred to as the Second Draft ISA. The Chartered CASAC approved the report on August 13, 2015. The CASAC's consensus responses to the agency's charge questions and the individual review comments from the CASAC Oxides of Nitrogen Review Panel are enclosed.

Overall, the Second Draft ISA is a much improved document and is very responsive to the CASAC's comments (EPA-CASAC-14-002, June 10, 2014) on the First Draft ISA. There are several recommendations for strengthening and improving the document highlighted below and detailed in the consensus responses. The CASAC believes that with these recommended changes, the document will serve as a scientifically sound foundation for the Nitrogen Dioxide (NO_2) National Ambient Air Quality Standards (NAAQS) review and no further review from the CASAC is needed.

The revisions made to the Executive Summary and Integrated Summary have markedly improved the document. Much of the technical language from the Executive Summary has been removed, improving the readability for the general public. The summary of major findings should be more concise. The revised Integrated Summary better synthesizes the key findings from each topic area and integrates the important information across the ISA that will be used to inform policy-relevant issues. The CASAC recommends that more clarification and justification be provided in the text for the changes made in causal determinations from the 2008 ISA, especially for the respiratory effects of long-term NO_2 exposure. The CASAC concurs with the EPA's causal determinations, but the reasoning behind the agency's decision to strengthen the causal determination for respiratory effects from long-term NO_2 exposure from "suggestive, but not sufficient, to infer a causal relationship" to "likely to be a causal relationship" needs to be more strongly justified and clearly articulated.

The revised chapter on atmospheric chemistry and ambient concentrations of oxides of nitrogen is substantially improved and responsive to the CASAC's previous recommendation to provide more detail on the spatial and temporal patterns in ambient NO_2 concentrations. For urban-scale variability, the use of the coefficient of divergence is appropriate and shows substantial differences in spatial patterns across different urban areas. The EPA should be clearer regarding the basis, including data sources, studies, averaging time, proximity to emission sources, and other relevant factors, of the reported concentration values in the summary table on near-road NO_2 gradients. Information on other near-source measurements (e.g., derived from proximity to marine ports, rail yards, airports, petroleum-related activities) would also be useful. The final ISA should include updated near-road monitoring data certified through the end of 2014. To the extent that sufficient data are available in the EPA's Air Quality System (AQS), it would be useful to extend the discussion of long-term trends back to 1980.

In response to the CASAC's previous comments, the 2nd draft ISA separates the discussion of exposure into a separate chapter. Overall, the discussion of exposure is markedly improved and does a good job of summarizing and contextualizing oxides of nitrogen exposure science within broader discussions of health effects. However, throughout the chapter the exposure terminology is inconsistent and, in some cases, confusing. The evaluation of exposure modeling should be revised. Statements promoting a specific modeling approach over another as being able to reflect spatiotemporal variability in oxides of nitrogen exposures are overstated. The discussion of spatial modeling (in particular, land use regression approaches) is incomplete and inaccurate in some sections. The EPA should give a more complete, accurate, and balanced presentation of modeling approaches, taking into account the limitations of each method. Greater attention should be given to physical activity as a modifier of exposure and to the role of time-activity.

The CASAC recommends that the evidence supporting changes to the causal determination status for oxides of nitrogen for associations with short-term exposures be based primarily on the findings from the controlled human exposure studies, as they alone are sufficient to justify the change.

The revised dosimetry and mode of action chapter is stronger, better-organized, and is very responsive to the previous CASAC comments. The addition of mode of action figures is helpful to better understand how the agency is synthesizing available information on pathophysiological mechanisms. However, the figures should be revised for improved clarity and accuracy.

The reorganization of the material in the ISA to integrate evidence across disciplines is a major improvement. The separate exposure chapter also facilitated the review of the health effects material in Chapters 5 and 6. The CASAC is impressed with the meta-analysis of controlled human exposure studies and finds that this analysis facilitates the inferences that can be drawn from the studies contained in the analysis. There should be greater discussion about how examination of the available data from the studies in the meta-analysis could aid in inferences about NO_2 health effects. In particular, there should be some examination of the relationship between ambient measures of NO_2 and personal exposures to relevant copollutants. Copollutants are generally well addressed in Chapters 5 and 6, but there should be a greater distinction between those copollutants of greatest concern and those of less concern in reviewing key studies.

The chapter on populations and lifestages potentially at increased risk for health effects related to NO_2 exposure is much improved and overall the conclusions are sound and well justified. The CASAC notes several areas for further improvement as detailed in the response to the charge questions.

The CASAC appreciates the opportunity to provide advice on the ISA and looks forward to the EPA's response.

Sincerely,

/S/
Dr. Ana Diez Roux, Chair
Clean Air Scientific Advisory Committee

/S/
Dr. H. Christopher Frey, Immediate Past Chair
Clean Air Scientific Advisory Committee

Enclosures

NOTICE

This report has been written as part of the activities of the EPA's Clean Air Scientific Advisory Committee (CASAC), a federal advisory committee independently chartered to provide extramural scientific information and advice to the Administrator and other officials of the EPA. The CASAC provides balanced, expert assessment of scientific matters related to issues and problems facing the agency. This report has not been reviewed for approval by the agency and, hence, the contents of this report do not represent the views and policies of the EPA, nor of other agencies within the Executive Branch of the federal government. In addition, any mention of trade names or commercial products does not constitute a recommendation for use. The CASAC reports are posted on the EPA website at: http://www.epa.gov/casac.

U.S. Environmental Protection Agency
Clean Air Scientific Advisory Committee
Oxides of Nitrogen Primary NAAQS Review Panel

CHAIR
Dr. Ana Diez Roux, Dean, School of Public Health, Drexel University, Philadelphia, PA

IMMEDIATE PAST CHAIR
Dr. H. Christopher Frey, Glenn E. Futrell Distinguished University Professor, Department of Civil, Construction and Environmental Engineering, College of Engineering, North Carolina State University, Raleigh, NC

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

Dr. Matthew Campen, Associate Professor, College of Pharmacy, University of New Mexico, Albuquerque, NM

Dr. Ronald Cohen, Professor, Chemistry, College of Chemistry, University of California, Berkeley, Berkeley, CA

Dr. Douglas Dockery, Professor and Chair, Department of Environmental Health, T.H. Chan School of Public Health, Harvard University, Boston, MA

Dr. Philip Fine, Deputy Executive Officer, Planning, Rule Development & Area Sources, South Coast Air Quality Management District, Diamond Bar, CA

Dr. Panos Georgopoulos, Professor, Environmental and Occupational Medicine, Robert Wood Johnson Medical School, Rutgers University, Piscataway, NJ

Dr. Jack Harkema, Distinguished University Professor, Department of Pathobiology and Diagnostic Investigation, College of Veterinary Medicine, Michigan State University, East Lansing, MI

Dr. Michael Jerrett, Professor and Chair, Department of Environmental Health Sciences, Fielding School of Public Health , University of California, Los Angeles, Los Angeles, CA

Dr. Joel Kaufman, Professor, Department of Environmental Health & Occupational Health, University of Washington, Seattle, WA

Dr. Michael T. Kleinman, Professor, Department of Medicine, Division of Occupational and Environmental Medicine, University of California, Irvine, Irvine, CA

Dr. Timothy V. Larson, Professor, Department of Civil and Environmental Engineering, University of Washington, Seattle, WA

Dr. Jeremy Sarnat, Associate Professor of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, GA

Dr. Richard Schlesinger, Associate Dean, Dyson College of Arts and Sciences, Pace University, New York, NY

Dr. Elizabeth A. (Lianne) Sheppard, Professor of Biostatistics and Professor and Assistant Chair of Environmental & Occupational Health Sciences, School of Public Health, University of Washington, Seattle, WA

Dr. Helen Suh, Professor, Department of Health Sciences, Bouve College of Health Sciences, Northeastern University, Boston, MA

Dr. Ronald Wyzga, Technical Executive, Air Quality Health and Risk, Electric Power Research Institute, Palo Alto, CA

Dr. Junfeng (Jim) Zhang, Professor of Global and Environmental Health, Division of Environmental Sciences & Policy, Nicholas School of the Environment & Duke Global Health Institute, Duke University, Durham, NC

SCIENCE ADVISORY BOARD STAFF

Mr. Aaron Yeow, Designated Federal Officer, U.S. Environmental Protection Agency, Science Advisory Board, Washington, DC

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

CHAIR
Dr. Ana Diez Roux, Dean, School of Public Health, Drexel University, Philadelphia, PA

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

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

Dr. Jack Harkema, Distinguished University Professor, Department of Pathobiology and Diagnostic Investigation, College of Veterinary Medicine, Michigan State University, East Lansing, MI

Dr. Helen Suh, Professor, Department of Health Sciences, Bouve College of Health Sciences, Northeastern University, Boston, MA

Dr. Kathleen Weathers, Senior Scientist, Cary Institute of Ecosystem Studies, Millbrook, NY

Dr. Ronald Wyzga, Technical Executive, Air Quality Health and Risk, Electric Power Research Institute, Palo Alto, CA

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

Consensus Responses to Charge Questions on
EPA's Integrated Science Assessment for Oxides of Nitrogen – Health Criteria
(Second External Review Draft – January 2015)

Executive Summary and Chapter 1 – Integrated Summary

The Executive Summary and Chapter 1 provide overviews of the ISA. The Executive Summary is intended to be a concise synopsis of key findings targeted to the broadest audience, whereas Chapter 1 is a more detailed synthesis of the ISA's most policy-relevant findings. The Executive Summary and Chapter 1 are revised to address the CASAC Oxides of Nitrogen Panel's advice to provide a more cohesive discussion of the array of issues that are considered in evaluating the causality of relationships between NO_2 *exposure and health effects. The revised discussions describe the extent to which available scientific information has addressed these issues and the uncertainties that remain.*

1. Please comment on how clearly the Executive Summary communicates the major findings of the ISA for a non-technical audience.

2. How well does Chapter 1 link together information about the distribution of NO_2 *in the atmosphere, exposure assessment, dosimetry, modes of action, and health effects to convey the major issues that need to be considered in evaluating scientific information on* NO_2 *exposure and health effects? To what extent does Section 1.4.3 address potential confounding factors?*

3. What are the Panel's views on how well Chapter 1 provides an integrated analysis of the weight of evidence for NO_2*-health effect relationships? For example, information on exposure assessment, dosimetry, modes of action, and health effects is incorporated into individual health effect discussions in Section 1.5 (e.g., respiratory effects, cardiovascular and related metabolic effects). Also, the section from the first draft ISA on confounding was removed and incorporated into each health effect discussion. To what extent is the causal framework transparently applied and the rationale for changes made (or not made) to causal determinations from the 2008 ISA for Oxides of Nitrogen clearly articulated in the Executive Summary, Chapter 1 and Table 1-1?*

The revisions made to the Executive Summary and Integrated Summary have markedly improved the document. Much of the technical language has been removed from the Executive Summary, improving the readability for the general public. In addition, there is more integration and less redundancy of the Integrative Summary in the Executive Summary, making the two sections more distinct. The revised Chapter 1 better synthesizes the key findings from each topic area, which are provided in the subsequent chapters, and integrates the important information across the ISA that will be used to inform policy-relevant issues. The tables on causal determinations and the figures on the suggested modes of action are very helpful and well done.

The summary of major findings at the end of the Executive Summary, however, is too wordy and could be more concisely crafted. Key findings related to the determination of causality could be further condensed (fewer findings and shorter text for each finding; delete subpoints).

The CASAC suggests that more clarification and justification be provided in Chapter 1 on the changes made in causal determinations, especially for the respiratory effects of long-term NO_2 exposure. The

reasoning behind the strengthening in causal determination from "suggestive, but not sufficient, to infer a causal relationship" to "likely to be a causal relationship" should be more clearly articulated, leaving no doubt about how and why this determination was made.

Chapter 2 – Atmospheric Chemistry and Ambient Concentrations of Oxides of Nitrogen

Revisions to Chapter 2 aim to address the CASAC Oxides of Nitrogen Panel's recommendation to describe in more detail spatial and temporal patterns in ambient NO_2 *concentrations and aim to clearly identify factors that may influence variation in exposure within the population and potential uncertainties in exposure estimates.*

1. Chapter 2 expands characterization of the spatial variability in NO_2 *concentrations within several U.S. cities (Section 2.5.2) and near-road gradients (Section 2.5.3) using information from U.S. monitoring networks and/or published studies. Please comment on the appropriateness of the content, interpretation, and scope of the material. How useful is the content and organization of Table 2-6, which synthesizes results from published studies of near-road gradients?*

The revised Chapter 2 is substantially improved. Sections 2.2 through 2.4 (Atmospheric Chemistry and Fate, Sources, and Measurement Methods) are appropriate in content and level of detail. The summary of national scale variability in section 2.5.1 is appropriate, but it should be noted here that, relative to the ozone and particulate matter (PM) monitoring networks, the NO_2 monitor density is significantly lower.

Table 2-4 provides summary information on both NO_2 and NO_x ($NO + NO_2$). For sites located near sources (e.g., any urban site), this is very useful information in the context of the potential for titration of NO to NO_2 to affect NO_2 gradients down-wind from the road. Pages 2-4 to 2-10 discuss NO_x and ozone chemistry in excessive detail; this material could be removed.

For urban-scale variability (Section 2.5.2), the use of the coefficient of divergence (COD) is appropriate, and shows substantial differences in spatial patterns across different urban areas. In the discussion of COD for Boston and Los Angeles (Figures 2-14 and 2-15), it should be noted that the difference in patterns of COD between these two urban areas (small for Los Angeles and large for Boston) is consistent with their very different spatial scales (Boston is much smaller in land surface area than Los Angeles).

Section 2.5.3, micro to neighborhood scale variability, receives the most emphasis given its importance. The discussion and presentation of near-road gradients in section 2.5.3.1 is useful. Table 2-6 summarizes NO_2 gradients from available research. This table combines results from a wide range of averaging intervals. Many of the studies used passive samplers with averaging times of days to weeks; the utility of these results is somewhat limited in the context of a 1-hour standard. Even for studies with short averaging times, the concentrations at the monitors nearest to the roadway are often surprisingly low (<50 ppb) except for the Los Angeles results. To enhance transparency and avoid misinterpretation, the EPA should be clearer about the basis of the reported concentration values in Table 2-6. The table should stand alone without having to look at the references. It is unclear if the data in Table 2-6 from studies with highly time-resolved data are long-term averages or some other metric. Some of the studies have a range of concentrations, while others have only a single value. Given the 1-hour NAAQS, a more detailed summary including 1-hour maximum concentrations and traffic information would be useful for

studies with time-resolved data, such as that presented in Figures 2-17 and 2-18 on pg. 2-58 and 59. Information on other near-source measurements (e.g., marine ports, rail yards, airports, petroleum-related activity such as hydraulic fracturing) would be useful, such as the rail yard studies noted on page 2-18. For improved clarity, EPA could consider splitting Table 2-6 into two tables with one focused on short term averaging times and the other focused on longer term averaging times.

2. Data from the U.S. near-road monitoring network became available after the first draft ISA, and the second draft ISA presents preliminary data for a small group of U.S. cities that had at least one full year of measurements. Please comment on utility to the review of the primary NO_2 *NAAQS of the presentation, interpretation, and scope of the discussion of the near-road network measurements.*

Section 2.5.3.2, Near-Road Monitoring, notes that the near-road data summarized in this draft (data through March 2014) are only a small subset of what is now or will soon be available as the network is built-out. The EPA released an update to the status of the near-road network build-out on May 20, 2015, posted on the Near-Road Monitoring web page at http://www.epa.gov/ttn/amtic/nearroad.html. This list of active sites includes meta-data about the sites and their target roads; it would be helpful to include information on overall traffic density (within 1000 meters) to provide context for gradient assessment. By July 1, 2014, 36 near-road NO_2 monitoring sites were operational. Out of an expected 75 phase 1 and 2 sites, 54 Near-Road sites are now operational. The final ISA should be updated to include certified data through the end of 2014. Summary tables such as Table 2-7 on page 2-60, which compares near-road NO_2 with area-wide NO_2 in the same urban area, are very useful and show that the 98th percentile of 1-hour daily maximum NO_2 concentrations in an urban area may not be at the near-road monitoring site. Tables such as these should be updated to include the more recent data.

Section 2.5.5 briefly discusses long-term NO_2 trends from the existing area-wide monitoring sites, going back to 1990. To the extent that sufficient data are available in AQS, it would be useful to extend the trend back to 1980. The Elizabeth Lab site at interchange 13 of the New Jersey Turnpike can be used to show trends of near-road NO_2. Although it is not classified a near-road site by EPA's criteria, it is representative of near-road NO_2 and has data going back to 1980 (see the trend plot in Mr. George Allen's comments). Sites other than those classified as near-road sites can be relevant to evaluating near-road air quality.

3. Section 2.5.3 further characterizes near-road NO_2 *concentrations with data that are available from networks outside the U.S. Data on near-road* NO_2 *were publicly available for several sites and years in London, U.K. but not Canada. To what extent are the statistics presented in Table 2-9 and the discussion of the London data useful and adequate for describing how monitor siting can affect characterization of the spatial and temporal patterns in* NO_2 *concentrations? Are the potential limitations (e.g., lack of traffic count data for roadside sites) of the London monitoring data appropriately described?*

The London near-road data presented in Tables 2-9a and b are useful despite the several differences noted in the text regarding fleet mix and other factors. The London roadside/curbside monitors listed are all within 10 meters of the road, closer than most of the U.S. near-road sites. Some of these sites such as Marylebone Rd. are very close to the road (~2 meters), are in street canyons, and might be considered on-road measurements. The curbside/street canyon monitor siting can help explain the large difference between the roadside and urban background concentrations. Traffic data for Marylebone Rd. is noted in

recent literature (80,000 vehicles per day). There is one new near-road site in Toronto, Canada (~400,000 annual average daily traffic), but it just started operation this year.

Chapter 3 – Exposure to Oxides of Nitrogen

As suggested by the CASAC Oxides of Nitrogen Panel, the discussion of exposure is separated into its own chapter and is considerably revised in response to the Panel's comments on the need for the discussion to better inform the interpretation of epidemiologic studies of various designs and exposure durations.

1. The exposure discussion is re-organized to clarify: a) the connection between particular exposure assessment methods and epidemiologic study designs, and b) the influence of exposure error on health effect associations from epidemiologic studies of specific designs. How explicitly and accurately is epidemiologic study design considered in the discussion of the utility and uncertainties of various exposure assessment methods, the nature of exposure measurement error, and the impact of exposure measurement error on NO_2*-health effect associations?*

Overall, the discussion of exposure in Chapter 3 is markedly improved and does a good job of summarizing and contextualizing oxides of nitrogen exposure science within broader discussions of health effects. However, the exposure terminology is inconsistent and, in some cases, confusing. Measurements at ambient sites, for example, sometimes are referred to as "exposures" and sometimes as "personal exposure measurements." The inconsistent terminology is particularly evident in the sections pertaining to long-term exposures.

The evaluation of exposure modeling should be revised. Statements promoting a specific modeling approach over another as being able to reflect spatiotemporal variability in oxides of nitrogen exposures are overstated. The discussion of spatial modeling (in particular, land use regression approaches) is incomplete and inaccurate in some sections. The EPA should give a more complete, accurate, and balanced presentation of modeling approaches, taking into account the limitations of each method.

The discussion in Chapter 3 does not consider physical activity as a modifier of exposure. In numerous studies, physical activity have been shown to be important determinants of oxides of nitrogen exposures. Similarly, greater attention to the role of time-activity should be considered, especially as it relates to the predictive power of land use regression (LUR) models (see Dr. Michael Jerrett's comments for greater detail).

Measurement error, and its different variants and components, is addressed in this chapter, but discussions of implications are not always clear. There is a lack of attention given to measurement error within cohort settings and its implications (see Dr. Lianne Sheppard's comments for greater detail). The form that error propagates from exposure to health effects estimates is complicated and, as discussed below, the CASAC recommends that the EPA consider adding a new descriptive table to clarify the potential impact of error on epidemiologic findings.

How effective is the discussion in facilitating the evaluation of the strength of inference from epidemiologic studies in Chapters 5 and 6?

The revised discussion of implications of exposure studies on epidemiology is much improved. In some of the studies presented, this discussion strengthens an interpretation of independent health effects attributable to NO_2. Despite this, the CASAC notes lingering difficulty within this chapter in ascertaining covariance between NO_2 and its copollutants in real-world settings. Although there is considerable effort dedicated to establishing differential NO_2-copollutant covariance patterns in non-ambient microenvironments, these results are inconclusive and, in some cases, contradictory (see Dr. Jeremy Sarnat's comments for greater detail). Given this residual uncertainty, the CASAC recommends that the primary evidence supporting changes to the causal determination status for oxides of nitrogen, particularly for associations with short-term exposures, be based on the findings from the controlled human exposure studies, rather than from the epidemiological studies. The controlled human exposure studies are unique in being able to isolate NO_2 exposure and response, in contrast to observational studies conducted in the real world.

To sort through the exposure results and their implications for epidemiological findings, it would be helpful to identify and evaluate attributes of exposure assessment approaches for informing causal determination. This could be done in the form of new descriptive table outlining specific exposure assignment methods, their sources of likely error, and value for informing broader measurement error questions and causal determination.

2. Section 3.4.4 expands discussion of the relationships of NO_2 *with copollutants and traffic noise for various short-term and long-term time periods as well as various exposure parameters (e.g., ambient, personal, indoor). To what extent is this information appropriately characterized and useful for the evaluation of potential confounding in epidemiologic studies in Chapters 5 and 6?*

The CASAC appreciates the greater attention to the relationships of NO_2 with copollutants within a heterogeneous mixture. The CASAC is not convinced that either the current science or the results in Chapter 3 offers a case for being able to adequately separate personal NO_2 exposure from exposure to traffic co-pollutants, including noise. One of the primary difficulties is an inability to apportion personal NO_2 exposure to fractions associated with indoor, outdoor, and traffic sources. Thus, it is possible that findings of weak correlations between total personal NO_2 and its copollutants may not rule out the possibility that personal exposures to NO_2 from traffic are still correlated with other traffic copollutants. Thus statements about potential confounding in Chapters 5 and 6 should, therefore be clear about how this potential bias has been considered.

Chapters 4 - Dosimetry and Modes of Action for Oxides of Nitrogen

Chapter 4 is revised to address the CASAC Panel's advice to improve characterization of the NO_2 *transport within the respiratory tract, existing dosimetric models, as well as mode of action for specific health outcome groups such as asthma exacerbation.*

The revised chapter is stronger, better-organized, and the EPA was very responsive to the previous CASAC comments. The addition of mode of action figures is helpful and justified, to better understand

how the agency is attempting to organize the disparate research information available into a pathophysiological mechanism.

1. The dosimetry section (Section 4.2) expands on the description of the epithelial lining fluid in the tracheobronchial and alveolar regions. Further, the deficiencies and uncertainties associated with the lack of a validated NO_2 *dosimetry model are more explicitly described. Please comment on the adequacy and clarity of these expanded discussions. To what extent does Section 4.2 address the reactive nature of* NO_2 *and its ability to pass beyond the epithelial lining fluid?*

The dosimetry section is much clearer and better organized. The chapter contains an appropriate level of detail. Answers to questions regarding the reactive intermediates are not available in the current scientific literature. It would be helpful to include the epithelial lining fluid (ELF) thickness in the regions shown so that Table 4-1 would make it clearer that penetration through the ELF is region-dependent. The bronchi and bronchioles are ciliated. Cilia may also be a target of NO_2. Some literature review on this may be warranted (see Dr. Michael Kleinman's comments for greater detail).

2. Section 4.3 discusses mode of action for specific outcome groups and also includes new figures that describe what scientific information is available on the key events and endpoints that make up the pathophysiological changes that lead to particular health effects. What are the Panel's views on the effectiveness of the organization around the outcomes of interest? To what extent do the new figures facilitate integration with the health effects evidence in Chapters 5 and 6?

The use of figures facilitates integration with the health effects evidence, although there are some questions related to the accuracy and layout of the figures. The caption legends could provide greater detail to better explain the figures. In Figure 4-1, the distinction between "Bronchoconstriction" and "Airway Hyperresponsiveness" is not clear (possibly this is related to specific versus non-specific tests of airway hyperresponsiveness, but this should be clarified). In Figure 4-3, "Vascular Activation" might be changed to "Endothelial Inflammatory Activation." Figure 1-2 is derived from these figures, but the use of white and grey backgrounds to emphasize the level of scientific certainty is a plus, and it is recommended that this format be used throughout. Additionally, it is suggested that common themes that link to varied outcomes (e.g., that neural activation may impact airway function and cardiac function) be noted.

Chapters 5 and 6 - Integrated Health Effects of Short-term and Long-term Exposure to Oxides of Nitrogen

In response to the CASAC Oxides of Nitrogen Panel's recommendations, the health effect evaluations in the second draft ISA more explicitly integrate various lines of scientific information and describe the strengths, sources of bias, and uncertainties in the evidence base. The revisions aim to address the Panel's comments on the need to more transparently apply the causal framework and clearly articulate the rationale for the causal determinations.

1. To more transparently characterize the weight of evidence for health effects, discussions are organized by specific outcome groups. For example, outcome groups under respiratory effects include asthma exacerbation and respiratory infection (versus respiratory-related hospital admissions). Within specific outcome groups, clinical outcomes and events are emphasized over subclinical effects that may

be more relevant to characterizing the mode of action. Please comment on the extent to which individual endpoints are appropriately placed into specific outcome groups. For example, how well does the discussion of asthma exacerbation integrate the evidence for relevant health endpoints across disciplines, including mode of action information? How clearly do the causal determinations identify the specific outcome groups that contribute most heavily to the conclusions?

The CASAC finds that the reorganization of the material in the ISA to integrate evidence across disciplines resulted in a major improvement. Tied to this, the separate exposure chapter also facilitated the review of the health effects material in Chapters 5 and 6. With respect to some of the clinical studies, the document should have greater discussion of "adversity" and "clinically significant" outcomes and how the results of these studies relate to these outcomes.

2. Section 5.2.2.1 expands discussion of an EPA meta-analysis of controlled human exposure studies of airway responsiveness in individuals with asthma. The methods for this meta-analysis are described in more detail, and additional analyses of individual-level data assess the magnitude and clinical relevance of effects. Further, sensitivity analyses are presented that demonstrate that the statistical significance, distribution of responses, and determination of clinical relevance are robust to the exclusion of full studies and the removal of repeated measurements. These analyses were recently published in Inhalation Toxicology in Brown (2015). Please comment on the extent to which the results from the meta-analysis, including the new analyses, are clearly described, appropriately interpreted, and informative to the evaluation of NO_2*-induced increases in airway responsiveness. Given that the results are now published in a peer-reviewed journal, what material that is presented in the manuscript could be removed from the ISA and referenced to the manuscript?*

The CASAC is impressed with the meta-analysis and finds that it facilitates the inferences that can be drawn from the studies included in the analysis. Given that the meta-analysis is published in a peer-reviewed journal, much of the detail in Chapter 5 could be deleted; references to the peer-reviewed publication would suffice. One disadvantage of retaining the detail for the studies considered in the meta-analysis is that it makes the discussion of other studies appear less comprehensive.

3. Drawing from Chapter 3, the health effect evaluations more critically evaluate the exposure assessment methods used in epidemiologic studies. Please comment on the adequacy and consistency with which exposure assessment, including the utility and uncertainties of the methods used and potential impact of exposure measurement error, is considered in describing the strength of inference from epidemiologic results. To what extent is available information on health effects related to personal and indoor NO_2 *adequately considered in conclusions?*

There should be greater discussion about how examination of the available data could aid in inferences about NO_2 health effects. In particular there should be some examination of the relationship between ambient measures of NO_2 and personal exposures to relevant copollutants. The discussion could make better use of differences in exposure associations by season in the discussion of results. More attention should also be given to studies that considered indoor exposures to NO_2 as the relationship between copollutants and NO_2 is likely very different than in studies which rely on ambient measures only.

4. Chapters 5 and 6 provide a more consistent critical evaluation of potential confounding by traffic-related exposures in epidemiologic studies. The potential for various copollutants, stress, and noise to confound NO_2 *associations with particular health effects is identified based on correlations with* NO_2

and similar health effects and mode of action (Section 1.4.3 and Table 5-1). Further, the strength of inference from copollutant models is assessed by considering the correlations reported between pollutants and potential for differential exposure measurement error. What are the Panel's views on the extent to which confounding by traffic-related copollutants and other exposures are appropriately and consistently evaluated?

The copollutant issue is generally well addressed. There should be more distinction between those copollutants of greatest concern and those of less concern in reviewing key studies. There can be more interpretation from studies of indoor exposure and for studies undertaken in different seasons. The indoor exposure studies can be informative because they do not have the same mix of copollutants as the outdoor exposure studies. More consideration of the modes of action associated with the various copollutants would also be of use. The agency may also want to revisit the selection of key studies/evidence in tables such as Tables 5-7 and 6-5. The McConnell et al. (2010) study should be given more attention in the discussion of key evidence.

5. The health effect evaluations describe in more detail judgments of the strength and limitations of the evidence, drawing upon information about study quality and evidence integration to form causal determinations (Section 5.1.2, Table 5-1). To what extent are the strengths, sources of bias, and uncertainties in the integrated evidence base adequately considered in forming causal determinations? How transparently is the causal framework applied to the evidence for each of the broad health effect categories in communicating the rationale for the causal determinations?

The rationale is generally well-articulated, although there is some concern about applying strict evaluation criteria to various studies. There is also some concern about the limitations of the classification scheme, especially with regard to carcinogenic impacts where the evidence is very mixed and limited, and where there is no discussion about co-exposures to known carcinogens that have been emitted jointly with oxides of nitrogen in the past.

Chapter 7 - Populations and Lifestages Potentially at Increased Risk for Health Effects Related to Nitrogen Dioxide Exposure

Chapter 7 is revised to address the CASAC Oxides of Nitrogen Panel's recommendation to provide a more integrated analysis of the weight of evidence for potential at-risk populations and lifestages and to expand the discussion of populations with proximity to roadways and risk of NO_2*-related health effects due to multiple co-occurring factors.*

1. The enhanced integrated analysis of at-risk populations and lifestages includes moving individual study results to tables and focusing the discussion on the synthesis of the health effects evidence as well as available information on exposure and dosimetry. Please comment on the effectiveness of the integrated analysis and the extent to which the strengths and limitations of the evidence are explicitly and consistently described in communicating the rationale for conclusions about at-risk populations and lifestages.

2. A new section (Section 7.5.6) describes what information is available on differences in NO_2 *exposure or risk of* NO_2*-related health effects for populations with proximity to roadways. To what extent does the added discussion accurately reflect the available information?*

The chapter is an improvement over the previous version. Overall, the conclusions in the chapter are sound and well-justified.

There are several areas that could be improved:

1) The overall framework and discussion of types of evidence is useful. It is not clear, however, why a framework for evaluating effect modification was not derived.
2) Most of the studies would not have had effect modification as the primary hypothesis. The document should acknowledge this limitation. It would be useful to identify which, if any, of the studies were undertaken with the effect modification as the primary hypothesis. The reason for concern is that such studies may have been under-powered to detect effects.
3) Each section should have some discussion of whether there are exposure differences versus differences in the health effects. The concern here is that if NO_2 is found to have causal associations with certain health outcomes, then higher exposures in a well-designed study likely would lead to higher effects. There is substantial evidence that groups in poverty or who are non-white experience higher exposures to NO_2, but the epidemiological evidence is still lacking. It is important to clearly show how the exposure differences follow socioeconomic status (SES) or racial gradients, because for those that are considered causal or likely to be causal, there is high potential for larger health effects even if the epidemiological evidence of a direct effect modification is lacking.
4) In some cases, such as older adults, there is a contradictory evidence base between epidemiology and controlled human studies. It is not as clear how the evidence is being weighted.
5) There were some instances of conflation between individual and group level SES. It is important to distinguish between these levels because they might have very different potential to modify the effects of NO_2 and could operate along difference biological pathways.

References

McConnell, R., Islam, T., Shankardass, K., Jerrett, M., Lurmann, F., Gilliland, F., Gauderman, J., Avol, E., Künzli, N., Yao, L., Peters, J., and Berhane, K. (2010). Childhood incident asthma and traffic-related air pollution at home and school. *Environmental Health Perspectives*, 118(7):1021-1026.

Appendix A

Individual Comments by CASAC Oxides of Nitrogen Primary NAAQS Review Panel Members on EPA's Integrated Science Assessment for Oxides of Nitrogen – Health Criteria (Second External Review Draft – January 2015)

Mr. George A. Allen

The second draft of Chapter 2 is substantially improved over the first draft. Sections 2.2 through 2.4 (Atmospheric Chemistry and Fate, Sources, and Measurement Methods) are appropriate in content and level of detail.

Section 2.5, Ambient Concentrations, is key to this revision due to the recent availability of NO_2 data (since the first draft) from some of the new near-road network monitoring sites. These new sites provide important additional information on both urban spatial variability from micro to neighborhood scales, and potential for exceedances of the 1-hour NO_2 NAAQS.

The summary of national scale variability in section 2.5.1 is appropriate. Table 2-4 provides summary information on NOx as well. For sites located near sources (e.g. any urban site), this is very useful information. Some of the minimum values seem odd however, such as an annual average NOx concentration of 0.1 ppb and NO of 0.01 ppb.

For urban scale variability (2.5.2), the use of the coefficient of divergence is appropriate, and shows substantial differences in spatial patterns across different urban areas.

Section 2.5.3, micro to neighborhood scale variability, appropriately gets the most analysis. The discussion and presentation of near-road gradients in section 2.5.3.1 is useful. Table 2-6 summarizes NO_2 gradients from available research. This table combines results from a wide range of averaging intervals. Many of the studies used passive samplers with days to weeks; the utility of these results is limited in the context of a 1-hour standard. It may help to separate results into 2 separate tables, one with averaging times > 6 hours and the other with 6 or fewer hours.

Even for studies with short averaging times the nearest concentration is often surprisingly low (< 50 ppb) except for the Los Angles results. It is unclear if the data in Table 2-6 from studies with highly time-resolved data are long-term averages or some other metric. Some of the studies have a range of concentrations, while others have only a single value. Given the 1-hour NAAQS, a more detailed summary would be useful for studies with time-resolved data, such as that presented in figure 2-17 and 2-18 on pg. 2-58 and 59.

A key point regarding near-far ranges/ratios is made on pg. 2-55, line 9. If the larger (mid to neighborhood) scale concentration is low (e.g., for less urban areas), that is likely to drive the strength of the gradient - not the concentration near-road.

Section 2.5.3.2, Near-Road Monitoring. As this section notes, the near-road data summarized in this draft (data through March 2014) is only a small subset of what is now or will soon be available as the network is built-out.

EPA/OAQPS released an update to the status of the near-road network build-out on May 20, 2015, posted on their Near-Road Monitoring web page at http://www.epa.gov/ttn/amtic/nearroad.html. This list of active sites includes meta-data about the sites and their target roads: http://www.epa.gov/ttn/amtic/files/nearroad/nearroadsites.xlsx.

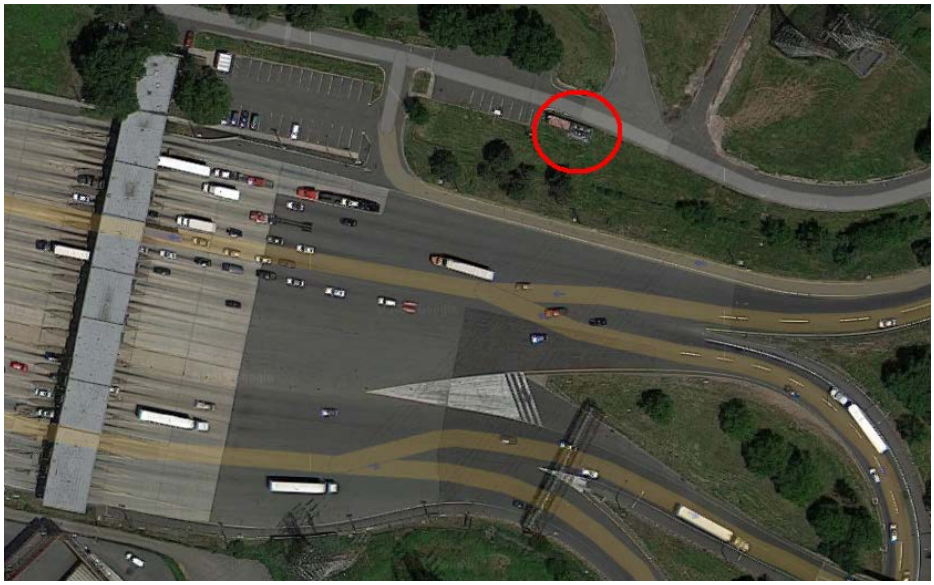
54 Near-Road sites are now operational out of an expected 75 phase 1 and 2 NO_2 sites. Two Phase 3 sites are operational, for a total of 56 sites. 36 NO_2 sites were operational by July 1, 2014.

Of the 70 sites with information on distance to the roadway, 11 are within 10 meters, and another 29 are within 20 meters. Only 9 are more than 30 meters from the road. Although the regulation requires sites to be within 50 meters, EPA has encouraged agencies to be within ~ 20 meters if possible, making the data more useful for estimation of curbside or on-road concentrations.

The final ISA should be updated to include as much of this additional data as possible, even if only for use in summary tables such as 2-7 on pg. 2-60 which compares near-road NO_2 with area wide NO_2 in the same urban area. This table is very useful and shows that the 98th percentile of 1-hour daily maximum NO_2 concentrations in an urban area is often not at the near-road monitoring site. It is worth noting that there are no exceedances of the 1-h NAAQS for the near-road data shown in tables 2-7 and 2-8.

Section 2.5.5 briefly discusses long-term NO_2 Trends. There is at least one long-term site that can be used to show trends of near-road NO_2. The Elizabeth Lab site at interchange 13 of the New Jersey Turnpike is not considered a near-road site by EPA's criteria; the NYC CBSA phase 1 near-road site is in Fort Lee, NJ. Although the Elizabeth Lab site does not technically meet EPA's near-road siting criteria, it is representative of near-road NO_2 and has data going back to 1980.

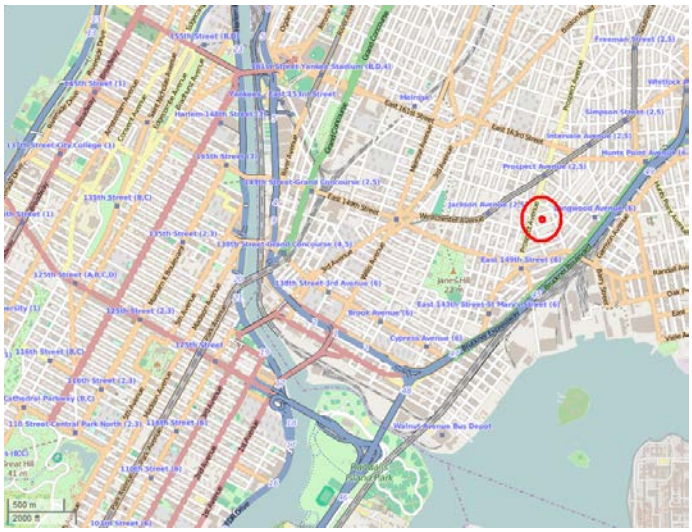
The Elizabeth Lab site location (circled):



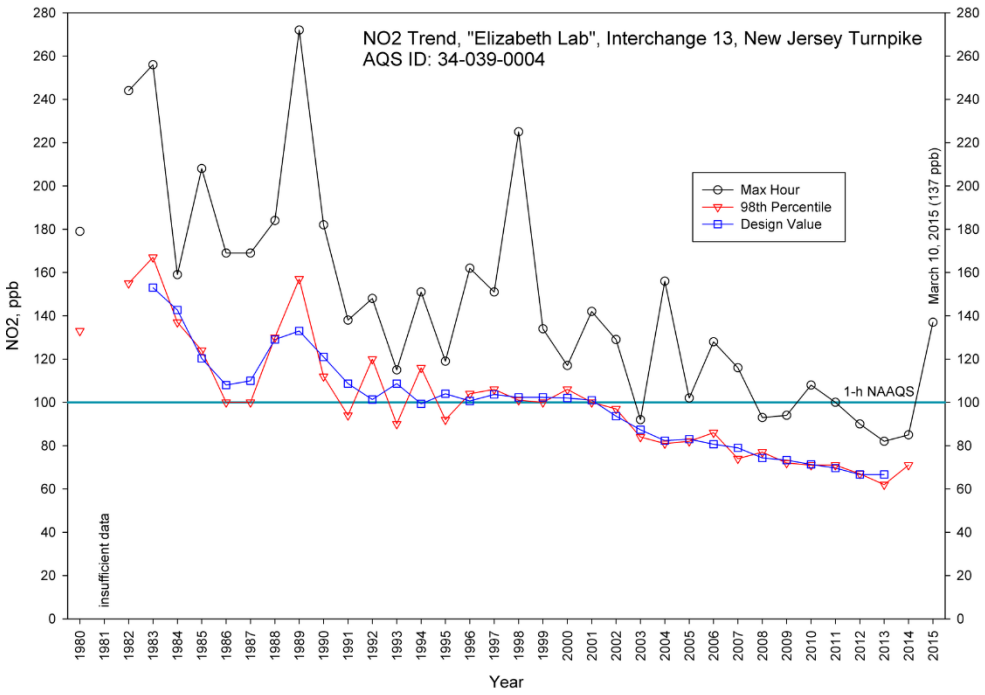
For context regarding the Elizabeth Lab site’s value for looking at long-term near-road NO_2 trends, a significant inversion/stagnation NO_2 event occurred March 10, 2015 in the NYC metro area that resulted in exceedances (> 100 ppb NO_2 1-hour average) at two sites. The Fort Lee near-road site did not have an exceedance during this event. A partial list of daily max NO_2 for this date at metro NYC sites follows.

Site	NO_2 (1-h max)	AQS ID	Notes
Elizabeth Lab	137 (NOx=618)	34-039-0004	also known as “Elizabeth Trailer”
NYC IS-52	122	36-005-0110	NATTS site
Jersey City	100	34-017-1002	not in AQS - urban canyon
Queens Coll. NYC	98	36-081-0124	NCore site
Ft. Lee NJ	81	34-003-0010	NYC phase 1 near-road site

The 2nd site with an exceedance on this day is a neighborhood-scale site (681 Kelly St.) at a school in the Bronx, circled below, approximately 420 meters northwest of I-278/895, the Bruckner Expressway.



The 34-year trend of hourly NO_2 for the Elizabeth Lab site shown below provides valuable context for near-road NO_2 concentrations. I encourage EPA to include this site in their near-road analysis.



The London near-road data presented in Tables 2-9a and b are useful despite the several differences noted in the text regarding fleet mix and other factors. The London roadside / kerbside monitors listed are all within 10 meters of the road, closer than most of the US near-road sites. Some of these sites such as Marylebone Rd. are very close to the road (~2 meters) and might be considered on-road measurements. This may explain the large difference between the roadside and urban background concentrations. There is some information on traffic count data for the London sites. The paper at http://www.atmos-chem-phys.net/11/6623/2011/acp-11-6623-2011.html has Marylebone Rd. traffic as 70,000 to 80,000 vehicles per day.

Charge question 3 for Chapter 2 notes the lack of any Canadian near-road NO_2 data. There is a traffic-oriented site run as part of the Southern Ontario Centre for Atmospheric Aerosol Research (SOCAAR) Field Measurement Facility in downtown Toronto, but this is not at a road with high AADT. A near-road site in Toronto has recently started at a location with 400,000 AADT, but data are not yet available.

Typo pg. 2-34, line 5: a word is missing?

Dr. Matthew Campen

The 2nd draft of chapter 4 represents a stronger, better organized chapter, and the authors were very responsive to the previous review. The addition of mode of action figures is helpful and justified, to better understand how the Agency is attempting to organize the disparate research information available into a pathophysiological mechanism.

Specific Charge Questions:

1. The dosimetry section (Section 4.2) expands on the description of the epithelial lining fluid in the tracheobronchial and alveolar regions. Further, the deficiencies and uncertainties associated with the lack of a validated NO_2 *dosimetry model are more explicitly described. Please comment on the adequacy and clarity of these expanded discussions. To what extent does Section 4.2 address the reactive nature of* NO_2 *and its ability to pass beyond the epithelial lining fluid?*

The revised draft is much clearer and better organized in the dosimetry section. The chapter contains an appropriate level of detail. What questions remain – regarding the reactive intermediates – are largely unknown.

2. Section 4.3 discusses mode of action for specific outcome groups and also includes new figures that describe what scientific information is available on the key events and endpoints that make up the pathophysiological changes that lead to particular health effects. What are the Panel's views on the effectiveness of the organization around the outcomes of interest? To what extent do the new figures facilitate integration with the health effects evidence in Chapters 5 and 6?

The figures are really quite nice for this integration. Additional details of pathways would be unjustified based on the current literature. In Figure 4-3, "Vascular Activation" might be changed to Endothelial Inflammatory Activation".

General comments:

Figure 2-19 and 2-20 do not reproduce well in grayscale. Consider changing some lines to dashed.

Page 4-32, nitrite is dismissed as a potential mediator of NO_2 toxicity, with justification from several therapeutic in vivo studies. I would recommend detailing specific NO_2/3 concentrations in serum and intracellularly, and noting the relative potential increase from inhaled NO_2, based on reports.

Eicosanoids are formed after NO_2 – are these due to enzymatic processes only, or as a reaction between NO_2 and arachidonic acid?

In the section on ANS (4-32 to 4-34), respiratory rate changes are used as a surrogate for neural activation. While there is certainly logic to this conclusion, it would be justified to note that many of the exposure studies did not specifically include permutations with a pharmacological inhibitor of the ANS, such as propranolol or hexamethonium. Without these, the respiratory rate changes are not mechanistically linked to ANS modulation. On 4-34, lines12-13, and appropriate statement is made

regarding atropine and vagal tone. Something similar should be noted earlier in the section for sympathetic activity.

Page 4-35, last paragraph, notes activation of NFkB and later about IL-6 and IL-8, but does not specifically note that these cytokines are under NFkB transcriptional regulation, which would tie the concepts together a bit better.

Pages 4-37 and 4-38, the alterations of selenium in the diet may alter glutathione, but was that measured and might other proteins be impacted by selenium availability? Just noting the limitation of the study may be worthwhile.

Dr. Douglas Dockery

Comments on Chapters 5 and 6 – Integrated Health Effects of Short-term and Long-term Exposure to Oxides of Nitrogen

1. *To more transparently characterize the weight of evidence for health effects, discussions are organized by specific outcome groups. Please comment on the extent to which individual endpoints are appropriately placed into specific outcome groups.*

   The organization of this ISA to truly integrate evidence across disciplines is a major improvement. With regard to the health effects chapters, bringing the evidence from toxicology, clinical studies, and epidemiology to bear on specific outcomes is much more in keeping with how we should be evaluating the scientific evidence. When these three approaches are considered separately, it is difficult for the reader to bridge to other chapters and to see how these independent approaches inform (or contradict) each other. This integrative approach presents the entire spectrum of information to the reader. When presented in separate chapters, there is the temptation to pay much more attention to the evidence from one approach, and only skim the other chapters. Moreover, in the approach specific presentation, there is a tendency to deconstruct the evidence from individual studies, rather than integrating across studies and approaches.

   Beyond the health chapters, I also see more integration with the other chapters. I see the Chemistry and Ambient Concentration (Chapter 2), Exposures (Chapter 3), and Dosimetry (Chapter 4) chapter as being more informative regarding health effects, and in this sense integrative also. Overall, I endorse and encourage this integrative organization.

2. *Please comment on the extent to which the results from the meta-analysis of controlled human exposure studies of airway responsiveness in individuals with asthma, are clearly described, appropriately interpreted, and informative to the evaluation of* NO_2*-induced increases in airway responsiveness.*

   I found the meta-analysis of controlled human exposure studies interesting and the data presented very informative. It is very appropriate to combine these studies in a meta-analysis. Given that this meta-analysis has been published, I agree that much of the detail is no longer needed in the ISA.

   With regard to the lack of a dose response in these data, I am not convinced that we would expect to see a dose response in these data. The combined data are for 116 NO_2 exposures among 72 individuals. Thus most individuals have only one exposure, which makes it hard to see individual dose response. Moreover, for these asthmatics, I would expect each to have a threshold for their individual response. Given a large enough sample, the distribution of individual response would show a continuous dose response. These controlled exposure studies are not measuring that threshold however, and 72 subjects is a small number to define a distribution.

Specific comments:

Page 5-21, Tables 5.2 and 5.3: Would be informative to include actual P-values rather than "n.s." if available.

Page 5-26, Tables 5.4, 5.5, 5.6: In describing table, it would be informative to specify that null is 0.50 (correct?). Also, please specify P-value rather than "n.s." In comparing "Exposure with Exercise" to "Exposure at Rest," should we compare to each other rather than to null? That is, is fraction responding with "exposure with exercise" significantly different from those with "exposure at rest"?

Page 5-29, Figures 5.1 and 5.2: Very informative. Could we see similar presentation for "exposure with exercise"?

Figure 5-2: What % of individuals had doubling? 19/66 = 23%

3. *To what extent is available information on health effects related to personal and indoor* NO_2 *adequately considered in conclusions?*

   Chapter 3 (Exposure Assessment) provides a very thoughtful and thorough examination of the state of knowledge regarding NO_2 exposures. Most importantly, Chapter 3 puts this information in context for evaluating the approaches and implications of these approaches for epidemiologic inference. Chapter 3 therefore provides an excellent base for evaluating exposure assessment methods used in the epidemiologic studies. This cross-chapter integration is a major strength of this revised ISA.

   The epidemiologic studies in Chapter 5 and 6 are evaluated in the context of the NO_2 exposure estimate approaches. Whenever possible, epidemiologic associations are compared stratified by the exposure assessment approaches. NO_2 exposure assessment is a rapidly evolving. This evaluation of the effect of alternative NO_2 exposure assessment approaches is a major strength of this ISA.

4. *Critical evaluation of potential confounding by traffic-related exposures in epidemiologic studies?*

   This ISA appropriately and consistently evaluates the potential for confounding by co-pollutants or other traffic related exposures. The ISA makes it clear that the failure to consider these potential traffic related co-pollutants or confounders is a weakness in almost all of the existing epidemiologic studies. This is not to say that these studies are not informative. Most importantly, the ISA appropriately integrates information across studies and approaches to evaluate the likelihood that such confounding is likely to be important in these studies.

5. *To what extent are the strengths, sources of bias, and uncertainties in the integrated evidence base adequately considered in forming causal determinations?*

   I can understand the need to provide guidelines for each of the three approaches - controlled human exposure, animal toxicology, and epidemiology, as in Section 5.1.2 and Table 5-1. As stated one would hope "to improve standards of reporting and ensure that data … can be fully evaluated." Indeed, it is helpful in integrating information across approaches to have some

understanding basic concepts of study design, subject/model selection, exposure assessment, outcome assessment, and statistical analytic approaches. Nevertheless, one should not expect such lists to substitute for experience and expertise in each of these fields of study.
Thus the criteria for evaluating epidemiologic studies (such as STROBE guidelines), should not be considered as defining study quality. These are guidelines representing common best practice. However, such guidelines are backward looking and would downweight innovative designs and analyses, or even deem them inappropriate based on historical practice. For example, time series or case-crossover studies would be classified as uninformative in NAAQS reviews in the 1990's using earlier sets of guidelines, whereas they have been shown to be robust, reproducible, and highly informative. Moreover, proscriptive guidelines are inconsistent with the integrative approach being used in this evaluation. Rather guidelines encourage a deconstructive evaluation of individual studies, that is checking off which study characteristics are not met within the list of positive characteristics.

There is no indication in the ISA that these guidelines are being used as criteria for acceptable studies. However, we should be on guard to ensure these do not become checklists in some sense.

Page 5-11: Not clear why confounding in epidemiologic studies is specifically highlighted by a separate Section.

Dr. Jack Harkema

Comments on the Executive Summary

General Comments:

The changes made to the first draft of the executive and integrative summaries (ES and IS, respectively) have markedly improved the document. Much of the technical language in the first draft has been removed from the ES improving the readability for the general public. In addition, there is more integration and less summary of the IS in the ES, making the two sections more distinctive and with less redundancy. The limited number of references and references only to specific sections, figures or tables are also appropriate.

The summary of major findings at the end of the ES, however, is too wordy and could be more concisely crafted. Key findings related to the determination of causality could be further condensed (less findings and shorter text for each finding; delete subpoints).

There still remains some inconsistencies between the ES and IS in the justification of causal determinations that need to be reconciled (see below under the IS review).

Specific comments:

Figure ES-1. This mode of action figure could be improved by indicating what key events or outcomes (health effects) that are the result of short- or long-term exposure. This should be articulated in the text as well.

lxxxii. lines 22-24. The statement, *The key evidence for an independent effect of* NO_2 *are the controlled human exposure findings for* NO_2*-induced increases in airway responsiveness* …, appears to contradict the statement in the IS, NO_2*-related decreases in lung function are observed in epidemiologic but not in controlled human exposure studies* (p.I-19, line 23-24).

lxxxii. lines 24-27. These two sentences do not make good pathobiological sense and should be revised. NO_2 reaction with antioxidants in lining fluid is thought to be a good thing (protection for the airway epithelium), but depletion of antioxidants and generation of reactive oxygen/nitrogen species could be detrimental (adverse effect) resulting in cell injury, airway dysfunction, and altered immunity.

lxxxv. lines 22. Authors need to clarify what they mean by *at low concentrations*.

Comments on Chapter 1 - Integrative Summary

General Comments

In general, the revised chapter better synthesizes the key findings for each topic area and integrates the important information across the ISA to inform policy-relevant issues.

The authors due a good job of justifying their recommendations for no causality change in specific areas, but some of the areas where they are recommending changes to the causality determination their rationale appears weak and needs more justification (e.g., respiratory effects associated with short-term exposure to NO_2; see specific comments below).

Specific Comments:

1-16. lines 1-2. Modes of action are proposed pathways that lead to an adverse health outcome(s) and are based on identified pathological key events. A mode of action helps to *explain* exposure/effect relationships but it does not *support* the relationship. The introductory paragraph for respiratory health effects is not well constructed and in some areas confusing. For example, if asthma exacerbation is the primary response to short-term exposure than the effects of NO_2 should be on the conducting airways and not on alveolar region of the lung (NO_2 uptake in the distal region of the lung is irrelevant and probably inconsequential) for this respiratory effect. It would be better to highlight what is known about NO_2-induced lung pathology that results in airway dysfunction related to asthma.

1-17 Figure 1-2. This expanded mode of action figure does illustrate key events and outcomes but their relationship to short- or long-term exposures are not illustrated.

1-17. line 5. Change *describe* to *suggest*. A mode of action pathway identifies potential key events in the development of potential health effects, but it is not definitive, like a discovered mechanism.

1-19. 28-37. The closing paragraph of this section suggests that no new data has been generated since the last review and that the justification for a causal relationship between short-term NO_2 exposure and respiratory effects is based on a better assessment/evaluation made by the authors. It is not clear how the authors arrived at this new insight. It appears to belittle any new scientific discoveries in this area since the last review that added to the weight of evidence. This closing paragraph should be more in line with the introductory paragraph of this section and the ES.

1-29-30. The potential confounding exposure factor of diesel engine exhaust particles (a known carcinogen) should be addressed in this section, especially since the authors are suggesting a change in the causal determination for NO_2 exposure and lung cancer.

Dr. Michael Jerrett

ISA – Summary Chapter

The document is much improved and more tractable. I had no major concerns, except with the characterization that most studies have not dealt with co-pollutants and this is only partly correct. Many studies have dealt with one or two co-pollutants, but not all; so this should be softened to haven't dealt adequately.

On types of evidence, there needs to be some mention of the scripted exposures such as the McCreanor et al. (2007) study. This is a hybrid between controlled exposures and epidemiology and could inform the determinations as well because the exposures are often more realistic than in chambers because they are in real world conditions.

Noise should be mentioned more prominently as a potential confounder.

The idea that we don't understand "double jeopardy" that well should be revised; there are specific studies on stress for example Shankardass et al. (2009) that do elucidate this area.

On exposures, could you delineate truck vs vehicle dominated roadways – this has important implications for understanding the impact of control measures given the high direct emissions of diesel?

There needs to be some caution in comparing to England because there is a larger proportion of diesel vehicles and therefore more primary NO_2. I wasn't sure why EPA didn't compare to the Canadian data. Health Canada has prepared all of this because they too are revising the NO_2 guidelines. Check with Barry Jessiman, Health Canada.

The discussion of roadway gradients could be refined by including mention of the types of roadways and volumes of traffic. In general much of the document talks about distance to roadways as though all roads are the same, whether they have 40,000 or 400,000 vpd. This needs to be addressed directly because it could explain why some cities observe higher NO_2 levels away from roadways.

Conceptual Model still Incomplete

- doesn't take physical activity into account
- fuzzy on what was meant at an indoor environment – in vehicle – how do we deal with that – not clear from the current presentation.
- The entire document is silent on intake fraction – this is a problem because so much of the NO_2 exposure probably occurs close to or at the source so the intake will likely be much greater

3.3 – Missed several studies on natural interventions – several physical activity studies (Andersen et al., 2015; Kubesch et al., 2015), road closures due to construction (Levy et al., 2004)

There is a literature on scripted exposure studies – the McCreanor study - increasingly important hybrid type models

3.3.1.1 – hard to compare between categories of distance to road with so many different impact zones cited

Could EPA conduct some systematic GIS analysis to show populations affected by near road exposures to make all the comparisons complete?

The review of LUR models is uneven and incomplete – many studies were missed (e.g. Jerrett et al. 2007 – and many more) – not sure why some were included and others not – needs to be clear; they did not deal at all with the issue of monitor siting and how this might affect the results; also there was no discussion of over-fitting and potential biases this introduces (see Beckerman et al. 2013a, 2013b - Beckerman et al. (2013b) only deals with PM2.5, but the point is that many of the current efforts in the literature on land use regression have over fit the data to the sample and probably are overestimating their predictive capacity. Beckerman et al. (2013a) does include NO2 also makes this point.) The comment that city-specific factors would differentiate the variable selected is partly correct, but how much would there be commonality – virtually all the models or the vast majority would have used some indicator of roadway proximity or traffic density. So there are some variables to go across many cities and locales.

Concerned about the spatiotemporal modeling discussion – not very sophisticated or subtle. See comments from Dr. Sheppard.

If you are going to mention other types of models that have not been used in Epi you should mention microsensors. I would recommend cutting those model types that have not been used so far.

On cancer, I found it odd that there wasn't any reference to IARC's recent pronouncements on traffic pollution being classified as a group 1 carcinogen; I also thought that the authors could improve the document by comparing, where possible, their conclusions to those of WHO or other national (e.g., Canada) or state level assessments (e.g. California).

Comments on Chapter 7

Please comment on the effectiveness of the integrated analysis and the extent to which the strengths and limitations of the evidence and the extent to which the strengths and limitations of the evidence are explicitly and consistently described in communicating the rationale for the conclusions about at-risk populations.

The chapter is an improvement over the previous version. Overall, the conclusions are the chapter are sound and well-justified.

There are several areas that could be improved:

1) The overall framework and discussion of types of evidence is useful. It was not clear, however, why a causal framework similar to earlier chapters was not derived.
2) Most of the studies would not have had effect modification as the primary hypothesis as effect modification. The authors at EPA should acknowledge this limitation. It would be useful to identify which, if any, of the studies was undertaken with the effect modification as the primary hypothesis. The reason for concern is whether most of the studies would have been under-powered to detect effects.
3) I would like to see for each section some discussion of whether there are exposure differences vs differences in the health effects. My concern here is that if we do find that NO_2 has causal associations with certain health outcomes, then higher exposures would, likely, in a well-designed study, lead to higher effects. There is substantial evidence the groups in poverty or who are non-white experience higher exposures to NO_2, but the epidemiological evidence is still lacking. But it is important to clearly show how the exposure differences follow SES or racial gradients, since for those that are considered causal or likely to be causal, there is high potential for larger health effects even if the epidemiological evidence of a direct effect modification is lacking.
4) In some cases, such as older adults, there is contradictory evidence base between epidemiology and controlled human studies. It is not as clear how the evidence is being weighted.
5) There were some instances of conflation between individual and group level SES. It is important to distinguish between these levels because they might have very different potential to modify the effects of NO_2 and could operate along difference biological pathways.

Also Arain et al. (2009) discusses temporal and land use controls over NO2 concentrations from Canadian data, which has more directly relevant traffic and industry impacted sites than in the U.S. - this might be helpful, since it more thoroughly combines meteorology data with land use influences on NO2.

References

Andersen, Z. J. J., de Nazelle, A., Mendez, M. A. A., Garcia-Aymerich, J., Hertel, O., Tjønneland, A., Overvad, K., Raaschou-Nielsen, O., and Nieuwenhuijsen, M. J. (2015). A study of the combined effects of physical activity and air pollution on mortality in elderly urban residents: the danish diet, cancer, and health cohort. *Environmental health perspectives*, 123(6):557-563.

Arain, M. A., Blair, R., Finkelstein, N., Brook, J., and Jerrett, M. (2009). Meteorological influences on the spatial and temporal variability of NO2 in toronto and hamilton. *The Canadian Geographer / Le Géographe canadien*, 53(2):165-190.

Beckerman, B. S., Jerrett, M., Martin, R. V., van Donkelaar, A., Ross, Z., and Burnett, R. T. (2013a). Application of the deletion/substitution/addition algorithm to selecting land use regression models for interpolating air pollution measurements in California. *Atmospheric Environment*, 77:172-177.

Beckerman, B. S., Jerrett, M., Serre, M., Martin, R. V., Lee, S.-J., van Donkelaar, A., Ross, Z., Su, J., and Burnett, R. T. (2013b). A hybrid approach to estimating national scale spatiotemporal variability of PM2.5 in the contiguous united states. *Environ. Sci. Technol.*, 47(13):7233-7241.

Jerrett, M., Arain, M. A., Kanaroglou, P., Beckerman, B., Crouse, D., Gilbert, N. L., Brook, J. R., Finkelstein, N., and Finkelstein, M. M. (2007). Modeling the intraurban variability of ambient traffic pollution in toronto, canada. *Journal of Toxicology and Environmental Health, Part A*, 70(3-4):200-212.

Kubesch, N. J., de Nazelle, A., Westerdahl, D., Martinez, D., Carrasco-Turigas, G., Bouso, L., Guerra, S., and Nieuwenhuijsen, M. J. (2015). Respiratory and inflammatory responses to short-term exposure to traffic-related air pollution with and without moderate physical activity. *Occupational and Environmental Medicine*, 72(4):284-293.

Levy, J. I., Baxter, L. K., and Clougherty, J. E. (2006). The air quality impacts of road closures associated with the 2004 Democratic National Convention in Boston. *Environmental Health*, 5(1):16+.

McCreanor, J., Cullinan, P., Nieuwenhuijsen, M. J., Stewart-Evans, J., Malliarou, E., Jarup, L., Harrington, R., Svartengren, M., Han, I.-K., Ohman-Strickland, P., Chung, K. F., and Zhang, J. (2007). Respiratory effects of exposure to diesel traffic in persons with asthma. *N Engl J Med*, 357(23):2348-2358.

Shankardass, K., McConnell, R., Jerrett, M., Milam, J., Richardson, J., and Berhane, K. (2009). Parental stress increases the effect of traffic-related air pollution on childhood asthma incidence. *Proceedings of the National Academy of Sciences of the United States of America*, 106(30):12406-12411.

Dr. Joel Kaufman

I am impressed with the amount of effort staff has put into improving the document since the first review draft. I appreciate that major sections have been re-organized in a way so that evidence can now be interpreted more clearly to contribute to the conclusions reached, particularly in chapters 5 and 6.

Section 5.2.2.1 expands discussion of the EPA meta-analysis of controlled human exposure studies of airway responsiveness. Now that the results have been published in the Brown et al paper (Inhalation Toxicology, 2015), the text in the ISA can presumably be dramatically reduced. Much of the text in the ISA appears nearly verbatim in the published paper now. I would suggest focusing on the paper's data summary and conclusions in the ISA. In addition, some of the text in the ISA (e.g., p 5-15, line 26 indicating "no meta-analysis") is no longer correct and needs to reflect the publication of this work. As in the prior draft, I continue to agree with the conclusion that an increase in nonspecific airway responsiveness in asthmatics is a valid and important outcome which serves as a valid surrogate for worsening of asthma. My confidence in this outcome and the findings in these studies is not materially reduced by the lack of NO_2 invoked response to specific allergen challenge nor the lack of response when exercise was part of the exposure regimen.

In the charge to the panel, we were also asked to determine the adequacy and consistency with which exposure assessment limitations (including methods, impact of measurement error, personal and indoor exposure concepts) were considered in describing the strength of inference from epidemiological studies in chapters 5 and 6. Overall, I consider the current draft to be quite good in this regards. These limitations are reasonably considered in evaluating individual studies and in forming summaries and conclusions. Further, I consider that difficulties in determining causal inference due to potential effects of correlated co-pollutants is consistently and clearly addressed in this draft.

Section 6.3.8 of the document discusses "Cardiometabolic Effects". While this is an increasingly popular term, in the context of the ISA I believe the evidence is better served to separate outcomes related to diabetes and related metabolic conditions from those that are more clearly cardiovascular diseases. It's true that obesity and diet can be common risk factors for diabetes and cardiovascular disease, and that diabetes is itself a risk factor for atherosclerotic cardiovascular disease. However, diabetes (and "pre-diabetic" conditions such as insulin resistance) represents a distinct set of pathological processes from heart disease and should be presented separately.

In chapter 7, staff has well-described the potential at-risk populations and lifestages for health effects related to exposure to oxides of nitrogen. Since the causal determinations related to these exposures are the most compelling for exacerbations of asthma and development of asthma, this should be the primary determinant of risk group identification. All asthmatics should be considered to be at risk of exacerbations of asthma from increased short-term exposure to high concentrations of nitrogen dioxide.

Although the document does review evidence related to genetic factors, my review of this indicates that none of the putative genetic factors can be considered to clearly represent an increased risk for exposure to oxides of nitrogen. While future research may demonstrate that there are genetic factors of interest, at this time the research is more informative of potential biological pathways than for identification of populations at risk.

Minor comments:

Chapter five discusses short-term exposure health effects of oxides of nitrogen. So while Table 5-1 is overall an extremely interesting and detailed table—which on balance adds value to the document—I found it odd that long-term exposure studies (i.e., studies for which spatial variation dominates exposure variation) are discussed in as much detail in this table. I would expect to see that in chapter 6. I do recognize that staff wished to have one comprehensive table and that it wasn't clear where to put it. One option, not clearly superior, would be to put the table in both chapters, but include the long-term relevant components only in chapter 6 and the short-term relevant components only in chapter 5.

I found the first paragraph in section 5.3.2 to be not written with sufficient precision. Not all epidemiological studies use diagnostic coding such as ICD systems (the highest quality epidemiological studies use individual record review for outcome assessment) so the discussion of nesting of ICD codes is not really appropriate for this introductory section. I think the authors are trying to get at something pathological rather than coding-related, but it isn't clear as written. Further in this paragraph there are more issues: ST segment depression is indeed a marker of myocardial ischemia—and in the right setting is actually pretty specific; however, the classical finding in myocardial infarction is ST segment elevation and not depression (now MIs—typically determined biochemically by blood markers of myocyte injury-- are sub-characterized as whether or not there is this characteristic elevation). A little clean-up of this paragraph would be helpful.

The wording in the sentence on line 9-10 of section 5.3.6.4 seems overly certain that "cross-sectional studies do not assess temporal relationships…" While there are limitations to interpretation and extra concerns about confounding, cross-sectional studies can sometimes be used to assess temporal relationships.
Sections 6.3.6 and 6.3.7 are a very heterogeneous set of outcomes. Some of the outcomes, such as surrogates for extent of atherosclerosis, are likely important and ultimately may be important in understanding the concentration-response relationship for long-term exposures. But many of these outcomes, such as heart rate variability or blood markers of inflammation and oxidative stress, are unlikely to elucidate much beyond potential mode of action. I don't think the document would be weakened by dropping these sections altogether.

Dr. Michael Kleinman

Comments on Chapter 4 – Dosimetry and Modes of Action for Oxides of Nitrogen

Chapter 4 is revised to address the CASAC Panel's advice to improve characterization of the NO_2 transport within the respiratory tract, existing dosimetric models, as well as mode of action for specific health outcome groups such as asthma exacerbation.

1. The dosimetry section (Section 4.2) expands on the description of the epithelial lining fluid in the tracheobronchial and alveolar regions. Further, the deficiencies and uncertainties associated with the lack of a validated NO_2 dosimetry model are more explicitly described. Please comment on the adequacy and clarity of these expanded discussions. To what extent does Section 4.2 address the reactive nature of NO_2 and its ability to pass beyond the epithelial lining fluid?

   The revised chapter provides a more thorough review of the basis of the model and identifies the alveolar region as the target zone in which the ELF is sufficiently thin that NO_2 and reaction/metabolic products can reach and interact with underlying tissue. It is noted that there are endogenous as well as exogenous sources for NOx and metabolites and that the concentrations of each of these are similar for many ambient situations. Table 4.1 clearly shows differences between human bronchoalveolar penetration and rat bronchoalveolar penetration. It would be helpful to include the ELF thickness in the regions shown so that the table would make it clearer that penetration through the ELF is region dependent. The bronchi and bronchioles are ciliated. The cilia are important components of the clearance process in the lung. Even though penetration to basal tissue is unlikely in the bronchial airways, the cilla may be affected and this can change rates of clearance. Reduced nasal mucocilliary clearance has been reported to be induced proportional to NO_2 exposure in urban motorcyclists[1]. The relationship of impaired clearance to increased susceptibility to respiratory system illnesses could be discussed and included in the mode of action discussion. While the evidence is still sketchy, there is an association between increased respiratory system infections and NO_2 exposures [2, 3]. There is also evidence of a relationship between increased exposure to NO_2 and ear infections [2].

2. *Section 4.3 discusses mode of action for specific outcome groups and also includes new figures that describe what scientific information is available on the key events and endpoints that make up the pathophysiological changes that lead to particular health effects. What are the Panel's views on the effectiveness of the organization around the outcomes of interest?*

   The figures provide a useful summary fo concepts. The figure captions could be phrased more specifically to highlight the key areas being discussed.

3. *To what extent do the new figures facilitate integration with the health effects evidence in Chapters 5 and 6?*

   They provide a useful structure for the discussion. It might be useful to somehow recapitulate these figures in Table 5.1 to provide a framework that helps focus the synthesis of information.

Specific Comments:

4-34 L 4 It could be noted that the diminished response seen after repeated NO_2 exposures is consistent with O3.
4-37 L 9 The comments on eosinophilic inflammation is linked to airway hyperreactivity later in the document, but could be mentioned here.
4-49 L 34 Alveolar bronchiolization due to NO_2 exposure is consistent with what has been reported for O3. Perhaps this suggests similarities in effects of oxidant gases in general.

1. Brant TC, Yoshida CT, Carvalho Tde S, Nicola ML, Martins JA, Braga LM, Oliveira RC, Leyton V, Andre CS, Saldiva PH *et al*: **Mucociliary clearance, airway inflammation and nasal symptoms in urban motorcyclists**. *Clinics (Sao Paulo)* 2014, **69**(12):867-870.
2. MacIntyre EA, Gehring U, Molter A, Fuertes E, Klumper C, Kramer U, Quass U, Hoffmann B, Gascon M, Brunekreef B *et al*: **Air pollution and respiratory infections during early childhood: an analysis of 10 European birth cohorts within the ESCAPE Project**. *Environmental Health Perspectives* 2014, **122**(1):107-113.
3. Aguilera I, Pedersen M, Garcia-Esteban R, Ballester F, Basterrechea M, Esplugues A, Fernandez-Somoano A, Lertxundi A, Tardon A, Sunyer J: **Early-life exposure to outdoor air pollution and respiratory health, ear infections, and eczema in infants from the INMA study**. *Environmental Health Perspectives* 2013, **121**(3):387-392.

Dr. Timothy Larson

Comments on the Executive Summary and Chapter 1

Communication to a non-technical audience

The Summary is generally well written and for the most part is understandable to an interested audience. Table ES-1 is an accurate summary of the ISA conclusions. Figure ES-1 is a good way to show the thinking on short term causality, although words such as "mast cell degranulation" and "epithelial barrier function" seem a bit too technical for a general audience.

The one section of the Summary that needs improvement is the "Summary of Major Findings". Many of the bulleted summaries of the major findings are awkwardly worded and frequently combine several concepts into a single bullet. It is hard to follow. In addition, the sub-bullets on page 89 (starting on line 1) contain too much detail for a non-technical audience.

Linking of major exposure issues relevant to causality

In general, this is a good summary of the important issues as they relate to causality. A few issues still remain. In Section 1.4.2 the reader is left with the impression that there is potentially a lot of exposure estimate error for long-term exposure estimates that vary over space. Further perspective on this uncertainty in the context of key studies would be helpful, especially given that long-term respiratory effects are likely causal. Most of the important long term epi studies cited in Table 6-5 have exposure estimates that are based on measurements taken near subject locations. Given the relative similarity in the results across these studies, it would seem that this measurement error is not as large as it theoretically could have been. Any context here would be helpful.

In Section 1.4.3 the issue of confounding is discussed in sufficient detail to follow the logic. The key issue of confounding by traffic related pollutants is identified. The discussion of the weight of evidence for each causal determination is easy to follow.

Of concern to this reader is the "suggestive" classification for NO_2 and lung cancer. It could just as reasonably kept its original classification of "inadequate" given that diesel engine emissions are a major source of not only NO_2 but also other co-pollutants and that diesel exhaust is classified as a known human carcinogen by IARC . It would seem that for this outcome the confounding issue for diesel exhaust is greater than that for general traffic related pollutants. Most of the studies listed in Table ES-1 and 6-20 either did not assess confounding in this way, or found no association with NO_2 in the presence of other co-pollutants such as UFP or BC known to be present in diesel exhaust. The rationale for the "suggestive" category needs more clarification. If one good epidemiological study (one of the criteria for this category) points to lung cancer, it should be identified. It is not clear which one that would be. From studies cited in the text and Table, Villeneuve et al 2014 measured NO_2, benzene, hydrocarbons but did not assess confounding by diesel exhaust. Raaschou-Nielsen 2010,2011 used a traffic marker NOx,

not NO_2. Jerret et al 2013 refers to NO_2 as a marker of traffic. Amigou et al 2011 looked at leukemia and used proximity metrics and traffic volume surrogates for exposure assessment, but did not measure NO_2 nor base their exposures on contemporaneous monitoring data; rather, they used a 4km2 smoothed spatial map to estimate background NO_2. Hystad refers to NO_2 as a traffic exposure indicator. Both Nyberg et al and

Nafstad et al used a traffic emissions model as input to a dispersion model validated with measurements of NO_2 to predict NO_2 at subject locations, but these same model predictions could also represent other traffic pollutants.

Specific Comments:

- P84 line 21 EC/BC, metals, or UFP are not obvious: one should not have to refer to the glossary
- P85 line 1 this wording implies that results from controlled human exposures of other traffic related pollutants were also summarized and considered. But that is not true. Improve the wording for clarification.
- P87 line 13 should "pattern" be "temporal pattern" ?
- P89 line 16 not true: some studies report null effect of BC
- P104 line 2 which studies of importance to long term effects as listed in Table 6-5 assessed the independent effects of residential proximity to roads? Asthma exacerbation or pulmonary function? It is certainly not the case for asthma incidence (p110, line 25). Needs clarification.
- P109 line 19 few epidemiological findings..
- P101 line 34 define "spatial misalignment" more specifically, presumably with subject locations vs LUR based monitoring sites
- P122, Table 1-1 the link to Tables 6-1 and 6-5 indicates that the confounding potential of traffic related pollutants in some of these studies is seemingly higher than others. The Vancouver cohort has much lower co-pollutant correlations than those from Gehring et al 2010. The summary statement in Table 6-5 is somewhat misleading in that regard. In the same Table 6-5, why do the findings of Shima et al 2002 have higher potential for exposure measurement error than the other studies? In that study, children went to school near both their home and their assigned monitor.

Comments on Chapter 3 – Exposure to Oxides of Nitrogen Measurements and Models

Page 3-4 line 30: Ross et al did not sample at all 150 sites simultaneously, only up to 25 in a given two week period. See also Allen et al Environ Res 2009 109(3), 334-342 for another extensive NO_2 passive sampling campaign.

Page 3-10 line 33: The model of Lindstrom et al 2013 is fundamentally different from that of Wilton et al. They are not directly comparable. Lindstrom predicted two week average measurements spatiotemporal framework that included static covariates and added the inherently spatiotemporal dispersion model as a separate term. Therefore the model even without the

dispersion predictions captures some temporal variability. Wilton simply added the dispersion model prediction as an additional covariate in the standard multi-linear LUR regression model so that the other covariates did not capture any temporal variability.

Page 3-11 line 21: Lindstrom's model (Lindstrom et al, 2013) should be discussed in either the Spatiotemporal Interpolation Model section or the Hybrid Models. It is a different category of model than the others discussed here. In addition to terms for the deterministic model predictions, it includes temporal basis functions that are combined with the static spatial covariates.

Page 3-18 line 29: OSPM predictions were compared with 2-week average NO_2 concentrations at > 200 sites at different times in 2006 in New York City with relatively low R2 =0.28 (Jensen et al Atmospheric Environment 43(2009) 4544-4556).

Page 3-19 Line 13 This discussion refers to hybrid dispersion models, not to be confused with other hybrid models that combine deterministic models with measurements. It might help to clarify this distinction in this entire section of the document. Maybe some number of subsections would help in this regard.

<u>Connections between exposure assessment method and study design:</u>

Table 3-1 is an important summary of the connection between a particular exposure assessment method and epidemiologic study designs. Below are specific comments regarding this important table.

- The table implies that the passive monitor method does not suffer from decreased correlation with distance from the monitor. Perhaps the Application column could be more specific, e.g., "short-term panel" could clarify a maximum spatial range of subject locations based on Figure 3-4.

- In this document, more weight is given to studies with measurements taken at subject residences than taken at central sites, but it is not clear if "short term panel" is referring to just personal sampling or also fixed site sampling at residences.
- The accompanying text also discusses the potential for exposure error if the passive samples are averaged over a week or more, even though this time scale is defined as "short-term".
- Are there examples of the CTM alone or the CTM-based Hybrid Method that show improved predictions of NO_2 with smaller scale CTM grid resolution? The Table implies this, but it is not clear if any of the cited studies actually show this. Perhaps the wording in the Table needs revision.
- The difference between the dispersion modeling and CTM methods needs to be clarified in the Table. Perhaps including "(e.g. Gaussian plume modeling)"
- The conclusions summarized in Table 3-1 regarding exposure errors should be clearly stated in the different subsections of 3.4.3.

Implications of exposure error for epidemiological studies

Community Time Series Study

The relevant literature addressing exposure error in time series studies by spatiotemporal interaction is discussed in detail. Table 3-13 summarizes the important results of Goldman et al (2012) who examined this issue in Atlanta. The text refers on line 11 on page 313 to a large negative exposure measurement bias of the area weighted average, but it must be referring to NOx not NO_2. The latter has a small negative bias according to the Table, a presumably important result given that the exposures are assigned to the total population. The original reference additionally cautions that thereare relatively few monitors for a given pollutant (5 for NO_2), and that these tend to be located in more heavily populated areas (three of them).

The study by Butland (2013) looked at modeled urban background NO_2 versus measured NO_2. The UK monitoring network emphasizes near-road monitors compared to the U.S. emphasis on monitors located further from the roadway. Thus the exposure assignment error is qualitatively different in this study than in the U.S. studies.

The interpretation by EPA of the study results from Dionisio et al 2014 regarding the bias introduced by variable indoor depositional loss seems reasonable but it was not discussed explicitly by the authors. Another factor that can decrease the correlation with outdoor monitors is the differences in sunlight and thus the NO_2/NOx ratio indoors versus outdoors.

Longitudinal Cohort Studies

The Szpiro and Paciorek (2013) paper seems important to this discussion of bias in chronic health effects studies. It would be useful to provide some context for those cases that are biased toward the null vs those that are biased away from the null. Specifically please clarify how their insights apply to long-term average NO_2?

Page 321, Line 17: This conclusion is important but it is stated awkwardly. Suggest eliminating the clause "..,such that the average total personal NO_2 exposure would necessarily be equal to or greater than the average personal exposure to ambient NO_2, .." for clarity.

Panel Studies

Page 322 Line 21 This concluding sentence needs further clarification. I interpret this statement as referring to the variation in the NOy concentrations (as well as the NOy/NOz ratios) across the urban area. Is this correct?

Confounding by Co-Pollutants and Noise

The discussion of potential confounding has been extensively improved from the last draft. It is an important part of the information relevant to the causality argument.
The Panel studies based on personal exposure measurements or outdoor residential measurements do not appear to have co-pollutant confounding, especially for subjects living far from busy roads. This result is not mentioned in the final summary paragraph on page 325. Given the extensive analysis of this issue in the main body of this Chapter, it deserves some concluding statement.

Dr. Jeremy Sarnat

General Comments

The Second External Draft of the NOx ISA is a coherent, extensive, and well-written report on the state-of-the-science regarding NOx health effects. EPA staff involved in preparing this draft deserve considerable credit for their clear responsiveness to the previous comments from the Review Panel and the public. The changes, both in structure and substance, are evident in the current version. The Second Draft ISA is transparent in addressing limitations, uncertainties, and methods used to inform causal determination.

There is notable, added attention in the current draft to issues of confounding and the assessment of NO_2 independent effects in both Chapters 3 and 5, which was lacking in the previous draft.

Chapter 1

- The strongest basis for causal determination in the association between short-term exposure and respiratory response are the few, controlled exposure studies involving airway responsiveness at environmentally realistic levels. These studies should, therefore, be presented first in this chapter, as they are in Chapter 5, as the rationale for proposing to strengthen causal determination status, rather than the additional results from observational studies which include NO_2 co-pollutants in multivariate model settings (See P1-17 through 1-19).

- While not specifically relevant for NOx, I question the designation of concentrations within two orders of magnitude from peak observed levels as being 'ambient relevant'. Even for controlled designs, studies of 5,000 ppb of NO_2, seem exceedingly high and not relevant to any realistic exposure scenario (Figure 3-1 on P3-22, clearly shows this). Even a one order of magnitude benchmark is high. Again, most of the science within this ISA involves concentrations/exposures far below the two orders of magnitude cutoff, but I think this ambient-relevant designation is worth reconsidering; especially as we move towards identifying potential risks associated with very low pollutants levels, including exposures well below the current NAAQS.

Chapter 3

- *Response to charge questions*: The Second Draft does a thorough job of summarizing and, more importantly, contextualizing NOx exposure science within broader discussions of health effects (Section 3.4), appropriate study design (e.g., Section 3.4.5), and measurement error and its implications (Section 3.4.3). The current structure of the NOx ISA represents a pronounced improvement over previous drafts.

- There is a critical assertion made within this chapter and repeated throughout the ISA, regarding observed correlation patterns between NO_2 and its copollutants, and their implications for assessing potential confounding. Specifically, instances of weak correlation between NO_2 and its copollutants, especially within indoor environments and as personal exposures, are presented as opportunities to disaggregate potential independent NO_2 effects. As stated in the Second Draft:

- "*[l]ow correlations between ambient* NO_2 *and personal measures of copollutants could support inferences regarding the independent effects of* NO_2." (P3-79)

For me, two main questions exist concerning observed weak correlations between indoor/personal NO_2 and its copollutants, including CO, UFP, EC, and VOCs.

- The first question is whether or not they are indeed real. As noted in the ISA, citing work by Meng et al. (2012) (P5-31), measurements of indoor and personal NO_2 [and here I would also include many of its copollutants] are frequently below detection and quantitated with increased analytical uncertainty, resulting in attenuated strengths of correlation between these non-ambient measurements and corresponding ambient NO_2 concentrations. For some of the studies reporting weak correlations, I suspect what we see are truly 'Type II-like' findings and that actual correlations between NO_2 and its copollutants, especially those from traffic sources, are likely stronger.

- A second question is whether appropriate correlation pairs are being examined. The Draft states that,
  - "*These observations [of weak correlations between personal* NO_2 *and its copollutants] provide further evidence that nonambient sources of* NO_2 *provide noise to the ambient* NO_2 *signal. At the same time, the weaker correlations between total personal* NO_2 *exposures and copollutant exposures indicate that for panel studies of total* NO_2 *exposure, ambient copollutants would be unlikely to confound health effect estimates for* NO_2 *exposure*." (P3-78).

  To me, the greatest source of uncertainty regarding causal inference is the specific role of NO_2 within a broader mixture of primary traffic emissions. To examine NO_2 as a potential marker for this mixture, therefore, then correlations should focus on associations between C_a *from traffic sources* and E_a *from traffic sources*. Since none of the exposure or measurement studies included within this ISA, even those where personal NO_2 exposures were conducted, are able to resolve source attribution on this level, this issue remains unanswered and central for defining the role of NO_2 in epidemiologic models, as either a contributing causal agent among a multiplicity of agents or a confounder (i.e., a non-causal surrogate of a true causal agent or mixture).

- Given the role of oxidative stress as a mediator of NO_2-related acute response, it would be useful to include a discussion of pollutant oxidative potential (OP) (i.e., its ability to generate reactive oxygen species), relative to its other copollutants. Clarifying NO_2 OP, particularly on a per mass basis, might be especially helpful addressing the plausibility of NO_2 independent effects at ambient-relevant levels.

- P3-20. Rate ratios from APEX models estimating associations between NOx and asthma ED were significant, but not significantly higher than corresponding rate ratios from models using alternative exposure assignment approaches.

- P3-56. In the Goldman et al (2010) paper, the authors report that, "instrument precision error increased with increasing concentration." It should be stressed that this trend is not consistent for all methods for measuring NOx or NO_2. Methods based on passive diffusion, for example, typically have greater precision error at low concentrations. This is relevant for the discussion of NO_2 correlation and potential condfounding.

- P3-83, Line 18. To be clear, differences between health effect estimates from models using C instead of E is not technically a form of epidemiologic bias. (A more accurate discussion of this concept can be found on P3-87, Lines 27 – 35.)

Chapter 5

- Much like Chapter 3, the revised chapter on health effects is much improved compared to the First Draft and is clearly responsive to the Panel's previous comments and discussions. I especially appreciate the careful attention given to the issue of potential copollutant confounding and the role of NO_2 within a traffic mixture, and am comfortable with most of the interpretations and conclusions made within this chapter (and within this ISA Draft, in general). The comments below mainly focus on one area of disagreement in my interpretation of the results and conclusion in the current Draft, relating to the recommendation to change the causal determination status of the association between short-term exposure and respiratory response.

- In general, much is made in the Second Draft of the consistency and coherence between the observational and controlled human studies for short-term exposure to NO_2 and respiratory response. Fundamentally, I still believe there are meaningful questions concerning the role of NO_2 as an indicator of a traffic pollution mixture. Although controlled exposure studies demonstrate the biological plausibility of independent, clinically-relevant NO_2 effects at stages along an asthma exacerbation pathway, I am not convinced of its primary and independent role in driving the corresponding responses within observational health effects literature. Based on this lingering uncertainty, I recommend that the decision to elevate the determination status to causal for short-term exposure and respiratory health be based primarily on evidence from the controlled exposure studies involving airway responsiveness.
  There seems to be some dissonance between conclusions in Chapter 3 and 5 regarding the potential for confounding of NO_2 independent effects. Throughout Chapter 5, short-term exposures to NO_2 are shown to be associated with respiratory morbidity 'alone and in combination with other pollutants' (P5-97). Chapter 3, however, seems clear in suggesting that it may be difficult to separate NO_2 exposures from general exposure to traffic pollution and corresponding health response:

  > "*Section 3.4.4 concludes that* NO_2 *concentration generally correlates spatially with other traffic-related pollutants in urban areas…With respect to exposure, these observations make it hard to distinguish* NO_2 *from other pollutants when considering the health impacts potentially attributable to each*." (P3-25)
  >
  > "*As a surrogate for traffic-related exposure,* NO_2 *concentration may do an adequate job of capturing spatial and temporal trends of traffic pollution*." (P3-25).

"*For traffic, NO (reacting to* NO_2*), CO, EC, UFP, and benzene are commonly coemitted and can be highly correlated with* NO_2 *in time and space*." (P3-97).

Uncertainty on this question is also related to statistical or modelling limitations. The main approach used in epidemiology and within the ISA to assess confounding, namely to model pollutants together within a co-pollutant or multivariate setting, is rightly acknowledged within the ISA to be lacking (P5-11, Line 20). Co-pollutant models are based on numerous assumptions, including linear associations among the independent terms, non-differential measurement errors among the copollutants, and other distributional assumptions that seem unlikely in most of the models conducted.

Taken together, the findings seem to point to the inability of observational designs, even those with excellent exposure assessment components, to conclusively disaggregate whether NO_2 is serving as a confounder or a marker of a source-specific mixture.

- P5-77. An example of a finding from a panel study that includes NO_2 and traffic-related VOCs that contradicts the interpretation of independent NO_2 effects is Greenwald et al. (2012). This paper showed that the outdoor and indoor BTEX VOCs were predictive of both increased pulmonary inflammation and decrements in lung function in a pediatric asthma cohort, where outdoor or indoor NO_2 concentrations were not significantly associated with either endpoint.

- The Second Draft presents limited results from alternative modeling approaches, in an attempt to consider whether NO_2 is acting multiplicatively, within a mixture, in eliciting acute respiratory response. Examples of these approaches are those in Gast et al. (2014), who used a C&RT approach and Winquist et al. (2014), who used a broad joint effects approach. The interpretation of finding from these studies is generally fair, although I'm not sure if either really strengthens inferences regarding NO_2 independent effects. Given the degree of concordance between the single-pollutant and joint effects models, for example, I would interpret the Winquist et al. (2014) paper as indicating NO_2 as a surrogate of a traffic pollutant mixture.

Dr. Richard Schlesinger

Comments on Chapter 4

Overall, the discussion of dosimetry in Section 4.2 is much improved, and the distinction between direct NO_2 reactivity in the lung and its (or its products) ability to pass beyond the ELF is well described. However, there is some room for further improvement, as noted below.

4.2.1. The information on ambient levels should be removed. There is an entire chapter that discusses this.

4.2.2

p. 4-3, line 6-10. While the ELF of the alveolar region clearly is involved in surface tension reduction, the ELF of the conducting airways is not. Thus, this sentence needs to be modified. Also, the sentence implies that the composition of the ELF throughout the lungs is similar, which it is not.

p. 4-4 Section Title Change title to "Reaction with Epithelial Lining Fluid" since the discussion goes beyond reaction merely with water.

p.4-3, line 20. Change "respiratory effects" to "toxic effects"

p. 4-4, lines 1-4. This is a bit confusing. The first sentence notes that the NO_2 dimer reacts with water, and then the next sentence notes that in aqueous solutions, NO_2 itself reacts with may solutes. This needs clarification. In fact, the entire paragraph should be rewritten as it tends to go around in circles, making the point less clear than it should be.

4.2.2.3

Some of the subsections have short summary paragraphs and others do not. It would help in evaluating the information if they all had a short conclusion at the end of each subsection.

4.2.2.4

p 4-20, line 4. Do the authors really mean O_3 or NO_2

The separation into distinct outcomes is good, but there still needs to be some additional "cleaning up" to avoid redundancy among the sections. There is the need to provide a summary for each outcome section that will allow the reader to reach some conclusion as to the role of NOx in affecting these outcomes, especially at ambient levels.

4.3.2

There is some redundancy in discussion of effects among the sections. For example, there are CV effects noted in section 4.3.2.2. and then again in 4.3.2.9, making it difficult to obtain some overall impression of the effect on the endpoint. Thus, some restructuring and consolidation of the section should be considered. In addition, some of the subsections have summary or conclusion sentences at the end while others do not; each section should have such a paragraph or sentence and also note whether the effects occur with ambient concentrations of NOx.

p.4-52, line 35. The term "spillover" here is not really scientific. Perhaps a better way to describe what is going on is "migration" of mediators from the lung into the circulation.

The figures provide somewhat of a summary of potential toxic pathways but otherwise do not really help with any discussions in the following chapters.

Figure 4.3 What is meant by vascular dysfunction as a result of vascular activation?

Dr. Elizabeth A. (Lianne) Sheppard

Chapter 3

General comments:
The reorganized text is a great improvement upon the previous version and I appreciate the efforts taken to be responsive to CASAC's concerns. I focused my attention on the long-term exposure and effects. I felt that while this chapter is a major advance, there are aspects that still need substantial improvement.

Organization

- Overall this chapter is much better organized and on target. Kudos to EPA staff! I focus most of the rest of my comments on suggested improvements rather than on the successes of this revision, of which there are many.
- There is text that appears under the wrong section heading. For instance, see page 3-19, the paragraphs starting lines 24 and 35.
- I'm surprised that in Section 3.4.5 "implication for epidemiologic studies" there is no attention to the new methods development for correcting for measurement error in cohort studies. The appropriate papers are (mostly) cited in the document, but there is absolutely no attention given to the methods development and the deeper understanding this brings to epidemiologic inference in cohort studies. Only the simulation studies and some of the definitions are reviewed. I suggest that under section 3.4.5.2 that a new subsection be added that covers the essence of the new methods developed by Spziro and others (specifically in papers published in 2011 that assume a parametric geostatistical exposure surface (note: the Epidemiology paper is cited but the Biostatistics one isn't)(Szpiro et al., 2011a; Szpiro et al., 2011b), and the Environmetrics paper published with Paciorek in 2013 that assumes a fixed exposure surface.)(Szpiro and Paciorek, 2013a, b) More generally, for all subsections of 3.4.5, I think the state of the methods for correcting for exposure measurement error should be the focus of the discussion of the subsection. Right now each subsection mostly focuses on comparisons of different estimate results, some are simulated and some are based on real datasets, though these distinctions aren't clear enough.

Clarity of writing: It will be essential that the text be edited for clarity, cohesion, and to ensure the appropriate points are being made. I have listed examples where corrections are needed in my detailed comments; this is not comprehensive. There are some confusing, unclear, overly general and/or misleading statements in this chapter. My written comments identify some of the sections that need attention. In particular Sections 3.2 and 3.4 need work.

Judgments and insights

- I'm finding it difficult to believe claims in the document that certain exposure estimation approaches are quite generally better than others, particularly when such statements are made without consideration of any additional information. For instance, as mentioned repeatedly in Chapter 6 (section 6.2), this document judges IDW and dispersion modeling to be more uncertain than e.g. LUR in their ability to represent the spatial variation in NO2, and thus by implication the reader is led to believe that these exposure estimates produce poorer health effect estimates. While there may be many examples of pairs of studies where this conclusion is valid, I could easily describe a pair of studies as a counterexample: one which uses LUR and the other which uses IDW to estimate exposure but where I would trust the IDW estimate more than the LUR estimate of long-term

average NO2. A few of the reasons would include the relative number of locations used in each study (more for the study using IDW), the representativeness of the monitoring locations (poorly aligned with study subjects in the study using LUR), the time period of measurement in each (much longer and better representing a full year in the study using IDW), the available covariates for the LUR (a limited or poorly chosen set would produce poorer estimates) and the authors' approach to model selection in the LUR (overfit LUR models will produce poor estimates of NO2, even when the resulting estimates are quite variable). It is not just the tools used to produce exposure estimates that matter, but the exposure study design and the application of the tools.

- The whole discussion of evaluation of models seems to be inadequately nuanced or informed by in-depth understanding of the prediction modeling methods. Early in the discussion of LUR models a footnote indicates "unless otherwise noted for the LUR studies, R2 refers to model fit." While not further defined, it appears that this means the R2 is equal to the square of the correlation coefficient from the regression of the predicted exposures on the monitored observations that are in the same dataset used to develop the LUR model. This kind of "in-sample" estimate is the last type of estimate of R2 I would wish to use to describe LUR model performance. It will tend to be too high, won't reflect overfitting of the LUR model, and won't actually inform us of the model performance we care about, namely predictive ability at subject locations. For this purpose, "out-of-sample" estimates of R2 are preferable; these are often computed using a technique called "cross-validation". Furthermore, out-of-sample R2 estimates can be obtained about the best-fit line (also called regression-based R2, computed as described above but using different measurement locations than the ones used to develop the model) or the 1:1 line (also called MSE-based R2). The latter R2 estimate gives a more complete picture of how the predictions at new locations compare to measurements. I suggest that the document warrants additional attention to how model evaluation is considered and discussed.

Major suggestions for improvement

- I suggest making a table summarizing all the studies discussed in Section 3.2. Easily being able to compare the models, evaluations, input data (including number of locations, number of time points (if relevant, i.e. more than 1), time scale), and results would help readers better understand this section. It is clear from the text describing the papers I am familiar with that the write-up is somewhat misleading or worse. Hopefully the addition of the table will help address this concern. The table should also indicate the time period and spatial domain of interest for each study.
- I suggest completely revamping the section on spatio-temporal modeling in Section 3.2 since this is becoming much more common in epidemiology (and it is the major approach used in EPA's MESA Air study). (This should replace/update the "spatiotemporal interpolation modeling" discussion starting on p 3-11.) Cite Lindstrom (2013)(Lindström et al., 2013), Szpiro (2010 Env)(Szpiro et al., 2010), Sampson (2011 AtEnv)(Sampson et al., 2011), Keller (2015 EHP)(Keller et al., 2015) and Li (2013 AtEnv; explicitly recognizing that Li is a fairly minor extension of the MESA Air spatio-temporal model)(Li et al., 2013). The spatio-temporal models discussed above use LUR, universal kriging (UK), and temporal trend function estimation using a singular value decomposition (SVD) of basis functions to capture both temporal and spatial variability. Many of their evaluations focused on spatial variation since that is the source of variation if interest for long-term epi studies. There are other papers that also report spatio-temporal models that might be cited in the spatio-temporal modeling section, even though they don't focus on NOx/NO2 applications. These include Paciorek (2009 AnnAppStat)(Paciorek et al., 2009), Yanosky, (2008 AtEnv; 2009 EHP)(Yanosky et al., 2008, 2009).

- Clarify the nuances of the R2's being reported in the document. (See my comments above.) It is not particularly helpful for readers to be comparing in-sample and out-of-sample R2 estimates across studies as though they are the same quantity. Furthermore, for out-of-sample estimates, there are additional distinctions to consider (see above). Precisely defining the R2's being reported throughout the document would help readers make fair "apples to apples" comparisons.

Response to charge question

1. The exposure discussion is re-organized to clarify: a) the connection between particular exposure assessment methods and epidemiologic study designs, and b) the influence of exposure error on health effect associations from epidemiologic studies of specific designs. How explicitly and accurately is epidemiologic study design considered in the discussion of the utility and uncertainties of various exposure assessment methods, the nature of exposure measurement error, and the impact of exposure measurement error on NO2-health effect associations? How effective is the discussion in facilitating the evaluation of the strength of inference from epidemiologic studies in Chapters 5 and 6?

The reorganized text is a great improvement upon the previous version and I appreciate the efforts taken to be responsive to CASAC's concerns. However, there are aspects that still need improvement. The measurement error discussion still needs to be refined and improved to actually focus on the methods advances as do some of the comments and generalizations about various exposure assessment strategies for application to epidemiology. (see my detailed comments above) EPA's goal of facilitating evaluation of the strength of inference from epidemiologic studies is not yet fully met. The text includes many details that make this chapter's material less useful than ideal for making judgments about the epidemiologic studies. Yet in other ways important details are missing (e.g., a synopsis of the statistical methods advancements on handling measurement error in cohort studies). The existing reviews of exposure assessment studies don't give the reader the deep insight needed to really understand their utility in epidemiologic study applications.

There are broad claims in the document that aren't always well supported. While I would agree that certain exposure estimation approaches should be better, claims are often made without good appreciation of essential additional information needed to evaluate any particular application. I'm finding it difficult to believe claims in the document that certain exposure estimation approaches are quite generally better than others, particularly when such statements are made without consideration of any additional information.

To meet the objective of facilitating the evaluation of the strength of inference from epidemiologic studies, I suggest EPA consider classifying exposure assessments for each epidemiologic study according to appropriateness for use in inference, using a system similar in spirit to those used for other judgments. For instance, each exposure assessment could be classified as strong, acceptable, weak, or inappropriate for the intended epidemiologic study. This would allow better judgment of epidemiologic studies based on the appropriateness of their exposure assessment. (A particular exposure assessment might get one judgment for one epidemiologic study and a different judgment for a different study.) The reasons behind the judgment should be provided as well.

Ultimately, given the current state of knowledge and the resources available, I think it will be difficult to successfully make all the changes needed to address this charge question and meet the objectives of this chapter.

Detailed comments for Chapter 3

- P 3-2 l 9: What does "research-grade" mean? How is this linked to central sites?
- P 3-7 paragraph starting line 4: This discussion is problematic. Please revisit.
- P 3-7 paragraph starting line 19: Ditto
- P 3-8 paragraph starting line 32: Isn't the point of the SA-LUR model to incorporate temporality into the model (through variables such as wind speed, etc)? So why does this paragraph open in a way that implies that the previous discussion wasn't about temporality?
- P 3-9 paragraph starting line 17: There are some misleading statements that should be corrected.
- P 3-10 line 23-5: While the statement is fine, I think the more important point is that for informing inference for epidemiological studies the comparison of the modeled estimates to measured values should be at locations that are relevant to the intended epidemiologic study.
- P 3-10: It is misleading to say Lindstrom 2013 "applied LUR" since the spatio-temporal model fit in that paper was much more complex than a LUR. Furthermore there were only two averaging times for model evaluation in that study: 2-week and long-term. There were no daily data; all the input data were on a 2-week time scale. The different model evaluation summaries (homes, snapshot, long-term averages at monitoring sites) each had different strengths and weaknesses and gave different insights into the spatial performance of the model.
- P 3-11 paragraph starting line 3: I suggest this paragraph discussing multiple linear regression as an "emerging exposure assessment method" should be dropped.
- P 3-11 discussion on spatiotemporal interpolation modeling needs to be updated and merged with the suggested new section I described above. Spatio-temporal modeling methods are no longer "emerging" for application to epidemiologic studies.
- P3-13 line 30-1: I don't understand the relevance to the ISA of CA DOT's lack of support for CALINE. Omit or clarify.
- P 3-16: Is Fuentes and Raftery the right reference for BME? And what about all the BME work in air pollution by Serre and his group?
- P 3-18 line 12-3: I wonder how many epidemiologic studies would care about model performance averaged over multiple locations?
- P 3-18 l 14: First these models should be defined.
- P 3-19 l 1: A stronger statement than "are not typically used" is necessary here. All probabilistic exposure models are inappropriate for use as exposures in epidemiologic studies because the probabilistic component induces measurement error in the health effect estimate. Probabilistic exposure models are very useful for risk assessment.
- Table 3-1 is a useful addition. Some details need to be corrected:
    - Revise the title: Most of the methods listed in the table are not "sampling methods".
    - Consider redesigning to have a set of columns for epidemiologic applications that rely on short-term exposure variation and another for those that rely on long-term averages
    - I don't understand the phrase "if the monitors are sited at fixed locations" under passive monitors. How else are passive monitors sited?
    - Kriging is omitted (ordinary and universal). Ordinary kriging could be considered with IDW.
    - The summary for spatiotemporal modeling reveals a lack of careful reading of the papers cited in the document. Please update. (See also comments above)
    - Parameterization modeling was never mentioned in the text

- P 3-23 l 13: The statement "exposure is likely to be underestimated" is too general to be true. The key challenge with IDW and other spatial smoothing methods is that with too few monitors the surface will be too smooth to capture the spatial variation of interest.
- P 3-56 Section 3.4.3.4: The focus of this section should be on the impact of instrument accuracy on epidemiologic study inference. How many studies differentiate exposure over a day? I would mention the broad classes of short-term time series and long-term cohort studies and talk about the role of instrument error in inference for each.
- P 3-57 Section 3.4.4: The overview of the confounding section is nice and could be a model for other sections (e.g. instrument error).
- The section heading for "3.4.4.3 Personal and Indoor Relationships between Nitrogen Dioxide and Copollutant Exposures" focuses on personal-ambient relationships.
- Section 3.4.5: This section needs work. Specific comments
    - Preface to Eq 3-12: Is this a model for cohort studies and a continuous outcome? This statement is so general to not be helpful. (We do use a model like this in our measurement error methods work where we make it clear the above two aspects are assumed.)
    - Lines 4-6: OK but I think this generality gets this statement into trouble. Usually we want exposure on the native scale for inference about health effects. The logit is a transformation of the outcome (for a glm) not a normalization; also there are details omitted here that are relevant to glms.
    - Line 7 "most …": This is much too strong a statement. What about survival outcomes or binary our count outcomes?
    - Paragraph starting line 9: Make sure to clearly distinguish pure vs. "-like" error and to define the latter. The definitions of "-like" error were developed for modeled exposures from e.g. a LUR. Also recent research has shown that there can be bias in either direction from Berkson-like error (see Szpiro & Paciorek 2013).
- P 3-83 l 27: Be precise. Were these pure or "-like" errors? I expect the former. Similarly, address statements on the following page on lines 2, 12-13, 13-14, 16-17. Some may be incorrect, or at least misleading as written.
- P 3-88 l 8-10: Strike this sentence. This work was based on simulation studies. The data were made up so the work could certainly be repeated for cohort studies. However the recent measurement error work for cohort studies makes important progress in a different way.
- P 3-38 l 4: If the central site monitor is truly systematically higher or lower, with no other missing features, then the slope (beta coefficient of interest) won't be affected.
- P 3-89 l 23-26: In this section these results deserve more comment since they are highly counter-intuitive to me. I'd like to know the details of what was done in the IDW vs. LUR to understand why this is true. For instance, if the IDW had the right time period, but the LUR didn't, this could affect the epi findings.
- P 3-90 paragraph starting line 33
    - l 33 "spatial errors": Be clear with terminology. Paciorek focused on confounding not error that is uncorrelated with exposure and outcome.
    - L 38: This reference to "effect of specification of spatial conditions" is unclear and misleading. Szpiro (2011)(Szpiro et al., 2011b) should not be reviewed in the same paragraph with Paciorek (2010 Stat Sci 25: 107-125) since the foci of the papers were entirely different!

    - The summary of Szpiro (2011) has numerous confusing, poorly worded, or erroneous statements. There was absolutely no confounding in the simulation study in that paper! (continues on following page)
- P 3-91:
    - Please also carefully review the discussions of Basagna and Szpiro & Paciorek for clarity and correctness. I had difficulties with both. In particular, with respect to S&P: This entire discussion completely ignores the new methods for correction that were developed and the assumptions that were made to accomplish this. One important role of this review would be to note that this approach is the wave of the future and that future epidemiologic cohort studies should be using measurement error correction methods since studies that don't do the correction get the wrong variance estimate for the health effect and may also need to correct for bias.
    - L 3: "predicting the true concentration" is not what is meant by the true exposure model in Szpiro (2011). The correctly specified model had all 3 covariates included while the misspecified model only included two of them.
    - L 5: Statement incorrect
    - L 6: Poorly worded
    - L 8-10: I don't understand this sentence.
    - L 11: Replace "of the correctly specified exposure model" with "in simulating the monitoring network data".
    - L 13: The bias in Szpiro (2011) was trivial and probably not worth mentioning in this discussion.
    - L 16: Insert "in the monitoring data" at the end of this sentence.
    - Paragraph starting line 36: This discussion needs to be fully revamped.
- P 3-92 l 17: In place of "criticized" it would be more correct to say "was pointed out to be a version of". Suggested new wording: "Spiegelman noted that the new measurement error correction methods developed by Szpiro & Paciorek (2013) are a version of regression calibration. This study ..."
- P 3-93 l 5: Will overestimating exposure *always* drive health effects towards the null? I can show a simple counterexample.
- P 3-93 paragraph starting line 14: Make sure the generalizations stated are correct and correctly qualified.
- P 3-95 l 25-7: This is too broad-brush of a statement. It needs to be qualified
- Section 3.5: There are some confusing, unclear, overly general and/or misleading statements in this section.

Chapters 5 & 6

I prioritized Chapter 6 and linkages with long-term exposures for my review.

I found the discussion of exposure assessment brought forward from Chapter 3 to be still in need of further refinement for clarity, accuracy, and utility for the purpose of judging the inferences that should be made from the epidemiologic studies. For example, text on pages 6-19 and 6-20 should be refined.

Response to Charge question

3. Drawing from Chapter 3, the health effect evaluations more critically evaluate the exposure assessment methods used in epidemiologic studies. Please comment on the adequacy and consistency with which exposure assessment, including the utility and uncertainties of the methods used and potential impact of exposure measurement error, is considered in describing the strength of inference from epidemiologic results. To what extent is available information on health effects related to personal and indoor NO2 adequately considered in conclusions?

I appreciate the inclusion of the exposure modeling approach in the summary figures and the exposure assessment details in the tables. The discussion of the utility and uncertainties of the methods used and the potential impact of exposure measurement error is less successful. For instance, the discussion in section 6.2.2.1 (starting p 6-19) needs some refinement.

The material in Chapter 3 did not focus on indoor and personal NO2 other than to look at their relationships with ambient NO2. Indoor and personal sources provide important information about health effects separate from ambient source NO2/NOx.

Detailed comments:

Please verify all the numbers in Table 6-1 reported for the McConnell (2010) study. I had a hard time finding all the HRs in the paper.

References

Keller, J. P., C. Olives, S. Y. Kim, L. Sheppard, P. D. Sampson, A. A. Szpiro, A. P. Oron, J. Lindstrom, S. Vedal, and J. D. Kaufman, 2015, A Unified Spatiotemporal Modeling Approach for Predicting Concentrations of Multiple Air Pollutants in the Multi-Ethnic Study of Atherosclerosis and Air Pollution: Environmental Health Perspectives, v. 123, p. 301-309.

Li, L. F., J. Wu, J. K. Ghosh, and B. Ritz, 2013, Estimating spatiotemporal variability of ambient air pollutant concentrations with a hierarchical model: Atmospheric Environment, v. 71, p. 54-63.

Lindström, J., A. Szpiro, P. Sampson, A. Oron, M. Richards, T. Larson, and L. Sheppard, 2013, A flexible spatio-temporal model for air pollution with spatial and spatio-temporal covariates: Environmental and Ecological Statistics, p. 1-23.

Paciorek, C. J., J. D. Yanosky, R. C. Puett, F. Laden, and H. H. Suh, 2009, Practical large-scale spatio-temporal modeling of particulate matter concentrations: Annals of Applied Statistics, v. 3, p. 370-397.

Sampson, P. D., A. A. Szpiro, L. Sheppard, J. Lindström, and J. D. Kaufman, 2011, Pragmatic estimation of a spatio-temporal air quality model with irregular monitoring data: Atmospheric Environment, v. 45, p. 6593-6606.

Szpiro, A. A., and C. J. Paciorek, 2013a, Measurement error in two-stage analyses, with application to air pollution epidemiology: Environmetrics, v. 24, p. 501-517.

Szpiro, A. A., and C. J. Paciorek, 2013b, Rejoinder: Environmetrics, v. 24, p. 531-536.

Szpiro, A. A., C. J. Paciorek, and L. Sheppard, 2011a, Does More Accurate Exposure Prediction Necessarily Improve Health Effect Estimates?: Epidemiology, v. 22, p. 680-685.

Szpiro, A. A., P. D. Sampson, L. Sheppard, T. Lumley, S. D. Adar, and J. D. Kaufman, 2010, Predicting intra-urban variation in air pollution concentrations with complex spatio-temporal dependencies: Environmetrics, v. 21, p. 606-631.

Szpiro, A. A., L. Sheppard, and T. Lumley, 2011b, Efficient measurement error correction with spatially misaligned data: Biostatistics, v. 12, p. 610-623.

Yanosky, J. D., C. J. Paciorek, and H. H. Suh, 2008, Predicting Chronic Fine and Coarse Particulate Exposures Using Spatio-Temporal Models for the Northeastern and Midwestern US: Epidemiology, v. 19, p. S211-S212.

Yanosky, J. D., C. J. Paciorek, and H. H. Suh, 2009, Predicting Chronic Fine and Coarse Particulate Exposures Using Spatiotemporal Models for the Northeastern and Midwestern United States: Environmental Health Perspectives, v. 117, p. 522-529.

Dr. Helen Suh

Comments on Chapters 3 and 7

Charge Question 3

The exposure discussion is re-organized to clarify: a) the connection between particular exposure assessment methods and epidemiologic study designs, and b) the influence of exposure error on health effect associations from epidemiologic studies of specific designs.

1. *How explicitly and accurately is epidemiologic study design considered in the discussion of the utility and uncertainties of various exposure assessment methods, the nature of exposure measurement error, and the impact of exposure measurement error on* NO_2*-health effect associations? How effective is the discussion in facilitating the evaluation of the strength of inference from epidemiologic studies in Chapters 5 and 6?*

2. *Section 3.4.4 expands discussion of the relationships of* NO_2 *with copollutants and traffic noise for various short-term and long-term time periods as well as various exposure parameters (e.g., ambient, personal, indoor). To what extent is this information appropriately characterized and useful for the evaluation of potential confounding in epidemiologic studies in Chapters 5 and 6?*

Chapter 3 provided a comprehensive discussion of exposures to nitrogen oxides, describing key issues related to the characterization of NO_2 exposures and their impact on our interpretation of epidemiological and other health studies. The Chapter is substantially improved over previous versions, with its new organization much better suited to not only describe our understanding of NO_2 exposures and factors affecting them but also to connect this understanding to help interpret epidemiological studies of NO_2, which it does successfully.

The chapter would benefit from additional relatively minor changes, as listed below.

- The Introduction (Section 3.1) describes the organization of the Chapter as including "methods to estimate personal exposure, current data used to characterize exposure to ambient oxides of nitrogen, exposure-related factors that influence interpretation of epidemiologic models of the health effects of oxides of nitrogen, and considerations for use of exposure metrics in epidemiologic studies of different design." These terms/phrases differ from section titles; for consistency, it would be helpful for the titles of the subsequent sections to have the same terminology.
- Figure 3-1 is a useful illustration of the variability in NO_2 exposures by location; in the discussion of this figure in the text and as a footnote, it should be mentioned that the data are based on different monitoring methods (in addition to exposure windows). This discussion (and figure footnotes) would be particularly helpful in light of the discussion regarding the method- and exposure window-specific considerations.
- Table 3-1 is a very helpful addition to the Chapter and helps to connect measurement and modeling methods to interpretation of health effect studies. However, the Table may be even more useful with the following changes or considerations:

- The errors and uncertainties section of the table focuses on method bias, but does not address method precision, which also has important implications for epidemiological studies and statistical power.
- It is not always the case that correlation between concentrations measured at central site monitors and "exposure" decrease with increasing distance, given the influence of roadways on NO_2 monitoring data and exposures. Also, it might be clearer to replace the term "exposure" with "outdoor locations", since exposure as used in the rest of the chapter can include factors other than outdoor location.
- Similarly, passive monitors do not always result in positive instrument bias, as negative biases have been demonstrated under stagnant conditions and using manufacturer recommended uptake rates.
- The term "exposure misclassification" often refers to exposure categories, while the methods included in the table all provide continuous measures of concentration or exposure. As a result, it may be more appropriate to refer to bias (rather than misclassification) in exposure estimation.

- Section 3.2.3 is a comprehensive and well-described review of measurement and modeling methods for NO_2. In the brief introduction to this section, it states that it will "outline various facets of characterizing NO_2 exposure", but should instead read "outline various facets of NO_2 measurement and estimation.
- Section 3.3 would benefit from a short introduction outlining the contents of the section (which tend to be relatively wide-ranging), prior to discussion NO_2 as an Indicator of Source-Based Mixtures. This introduction would help to provide a road-map for the section and to explain the purpose of the discussion, particularly as it relates to Section 3.4.
- The section on confounding (Section 3.4.4) should be mention that the potential for confounding of NO_2 impacts by co-pollutants can vary by the health endpoint of interest.
- Page 3-74, line 3-5. The sentence beginning "The next section…" seems out of place. Some clarification or re-wording is needed.
- Much of the discussion of results for Community Time-Series Studies (Section 3.4.5.1) are based on data from Atlanta. It is not clear that findings from Atlanta are generalizable to other cities in the US, which should be stated, with perhaps pointing to the need for further study in other locations.
- Section 3.4.4 was useful and appropriately characterized.
- In addition to the conclusion section at the end of the Chapter, it may be helpful for the reader to include a brief conclusion paragraph after each major Chapter section.

Charge Question 7

Chapter 7 is revised to address the CASAC Oxides of Nitrogen Panel's recommendation to provide a more integrated analysis of the weight of evidence for potential at-risk populations and lifestages and to expand the discussion of populations with proximity to roadways and risk of NO_2*-related health effects due to multiple co-occurring factors.*

1. *The enhanced integrated analysis of at-risk populations and lifestages includes moving individual study results to tables and focusing the discussion on the synthesis of the health effects evidence as well as available information on exposure and dosimetry. Please comment on the effectiveness of the integrated analysis and the extent to which the strengths and limitations of the evidence are explicitly and consistently described in communicating the rationale for conclusions about at-risk populations and lifestages.*

The Chapter is well-written, -organized, and -reasoned, providing a solid scientifically-based rationale for its conclusions. The movement of study results to tables was a welcome change and allowed for a streamlined and as a result, more useful and thoughtful Chapter. The Chapter would benefit further from greater reference to topics and issues raised in earlier ISA chapters.

Other comments include:

- No data are provided for COPD in Table 7-2. It may make sense to remove the row for COPD and instead include COPD in the footnote and note that it is comprised of chronic bronchitis and emphysema. Else, it gives a false impression that no individuals have COPD within the US population.
- While not relevant for many potential at-risk sub-groups with pre-existing diseases, it is likely important to include a discussion of exposure-related issues for people with asthma (even though it is included later in the section discussing children). For example, both children and adults with asthma may have different time/activity patterns as compared to other groups or may stay indoors on days with high air pollution levels, possibly reducing the ability of studies based on central site NO_2 concentrations to detect NO_2-related health impacts.
- The lack of evidence showing different time-activity patterns for older as compared to younger ages, the limited evidence from controlled human exposure studies, and the well-documented impacts of PM2.5 and ozone on hospital admissions and other health endpoints in older adults suggests the possibility of confounding of NO_2-impacts in older adults by $PM_{2.5}$, ozone, or other correlated co-pollutants. Some discussion of this possibility should be provided.

2. *A new section (Section 7.5.6) describes what information is available on differences in* NO_2 *exposure or risk of* NO_2*-related health effects for populations with proximity to roadways. To what extent does the added discussion accurately reflect the available information?*

This new section focusing on populations living near roadways and spending time near traffic is appropriate and is important addition to this Chapter. It is well-written, but needs greater background/review at the beginning of the section to link roadway proximity or time spent in traffic to elevated NO_2 concentrations. For example, the paragraph on page 7-52, lines 1-5 states that high NO_2 concentrations are found within 20 m of roads, while the following paragraph (as well as other paragraphs, such as on page 7-55, lines 10-26) presents data for population living within 100-250 m from roadways. As a result, these paragraphs are spatially inconsistent and together with the subsequent paragraph (page 7-53, lines 6-19) suggest that most of the US population is not exposed to elevated NO_2 exposures from busy roadways.

Dr. Ronald Wyzga

Charge Questions 1-3:

By and large these sections reflect the overall policy of the Agency and the content of the remaining Chapters. In several areas a little more detail or clarification would be helpful. See my detailed comments below.

Detailed comments on Preface and Chapter 1.

p. xlvii: ll. 17-19: Information on mechanisms can aid in the interpretation of these results.
ll. 29-33: It should be pointed out that there are generally fewer degrees of freedom on cross-sectional study, wh9ich makes consideration of an extensive set of confounders difficult

p. lii: Table 1 - Consistency: I'm not sure what is meant by the sentence: "Elevated risks are not defined by statistical significance." This sentence need be clarified.
Strength of the observed association: "may or may not"

p. liii: ll. 6-12: but statistical significance is nevertheless informative and should be indicated.
l. 33: insert "can" before represent.

p. lv: Table II - Causal relationship: Is two orders of magnitude too high? Some discussion would be welcome.

p. lix, l. 4: or a different threshold.

p. lxi, ll. 5-9: The ATS definition also has a statement about the concurrent occurrence of symptoms. This discussion need be modified.
l. 15: change would to could.
The bottom line is that there is no clear definition of what is adverse. It reflects considerable judgment.

p. lvi, l. 30: is two too high? Discussion please.

p. lxxvii, ll. 16-21: It should be clarified that many emissions are of NO which converts to NO_2. This can impact the gradient with respect to difference from roadside emissions.

p. lxxix, ll. 20-33: This issue is complicated by the simultaneous presence of confounding by other pollutants. This need be discussed.
l. 38: The other traffic-related pollutants should be indicated.

p. lxx, ll. 14-19: Averaging time need be stated. Also some mention of current levels would also be helpful here.

p. lxxxiii: figure ES-1: should "asthma attack" be replaced by "asthmatic response"?

p. lxxxvii, ll. 4-7: Error can also change the shape of the dose-response function.

p. 1-01, ll. 15-33: I urge the Agency to present any available results from the near-road network. Results for the contemporary US would be particularly valuable.

p. 1-11, ll. 19-30: It would be worth mentioning that indoor sources of NOx can be important and influence personal exposures.

p. 1-12: The potential emissions of NO and their conversion to NO_2 should be mentioned as influencing the results.

p. 1-13, l. 16: Are these correlations for personal or for ambient measures?

p. 1-18, ll. 14-20: The most important co-pollutants to consider should be highlighted.

p. 1-23, ll. 12-17: See above comment.

p. 1-37, ll. 27-29: The high correlation between 1-h max and 24-hr ave Nos should be noted.

p. 1-41, l. 8: may or may not; on-road exposures, if high, could not be reflected in these averages.
ll. 31-38: The correlations between NO_2 and co-pollutants may differ by concentration level (and place of measurement).

p. 1-44, l. 10: How does EPA interpret this definition?
ll. 12-17: I believe the ATS definition also mentions the co-occurrence of symptoms. This should be stated.

Comments on Chapter 2

Charge Questions 1-3: I would urge the Agency to present all available on-road measurements that are currently available. These data should then be contrasted with the data from London.

Detailed comments:

p. 2-71, Figure 2-21: Given the concerns about long-term effects, including cancer, it would be important to have some indication of trends in NO_2 levels from 1975. Any information about these trends should be included.

Comments on Chapter 3

Charge questions 1-2: I appreciate the detailed information presented here, and I compliment the Agency for its organization of this extensive material. There should be greater discussion about the implications of the relatively weak correlations between ambient and personal NO_2 measures. The differences by season would be highlighted and revisited again when the epidemiology results are presented.

Detailed comments:

p. 3-25, l. 34: When the on-road measurements become available this should be updated.

p. 3-26, Figure 3-2: what is the difference between VOCD1 and VOC2?

p. 3-31, Table 3-3: Define what is meant by Reference Site and by Regulatory Site? Should the reference be Matte et al or Ross et al (2013)?

p. 3-47, Table 3-5: The Personal-Ambient Slope for Sahsuvaroglu results are strange. Why is the total so much more highly correlated that the temporal subsets.

p. 3-49, Table 3-6: clarify the difference between outdoor and ambient monitors?

p. 3-50, ll. 2-7: This has serious implications for epidemiology studies.

p. 3-52, Table 3-7: The averaging time is not clear. Ideally results should be presented for both hourly and annual concentrations given that the NAAQS utilize these averaging times.

p. 3-83, l. 32: Is there a sign missing for NOx?

p. 3-84, ll. 19-22: but if the correlation between ambient and personal measures of NO_2 is very low, we have a greater problem than measurement error.

p. 3-85, ll. 8-11: See above comment.

Comments on Chapters 5 and 6

Charge Questions:

1. The organization is helpful. My concern is that greater weight should be given to analyses that considered relevant co-pollutants and studies that considered personal (and possibly indoor) exposures. The text need to explain why such studies need to be emphasized; it also tends to ignore the relatively low correlations between personal and ambient exposures presented in Chapter 3. Chapter 5 does not to my mind adequately summarize the results from human clinical studies. The results are often poorly summarized; moreover emphasis should be given to those studies with exposures near contemporary ambient levels. There should be greater integration of results from clinical studies with alternative definitions of “adverse” or “clinically significance”.
2. I need to carefully read the Brown (2015) paper before commenting. By and large the document does a good job of summarizing and presenting the results.
3. I believe that Chapters 5 and 6 need to do a better job in utilizing exposure information presented in Chapter 3. Ambient-personal correlations are generally small – so small that exposure misclassification is likely of greater importance than exposure error. There should be some discussion of this issue in the Chapter. I believe that this small correlation, as well as seasonal differences in this association, are not adequately addressed in the conclusions.

4. The co-pollutant issue is certainly well-addressed with two exceptions. There needs to be more emphasis given to studies that consider this issue. This is not always the case. Also, clearer distinction must be made between the more relevant co-pollutants of concern and those of lessor concern. Seasonal differences could also play a role here.
5. The rationale are reasonably laid out. I think the discreteness of categories bothers me. So often we see the terms "limited"" and "inconsistent". These generally translate into "suggestive" because the only alternative appears to be "inadequate". I frankly would like to see a category between these two because so many of the endpoints considered fall in between these two. I am particularly troubled by the carcinogenicity classification.

Detailed Comments:

Section 5.2.2.1, p. 5-15: how important is the dose of the challenge type? I frankly don't know, but some discussion of this would be helpful.

p. 5-23, Table 5-3, Tiedl et al 2012: is this results protective?

p. 5-34, l. 8: How was "clinically relevant" determined?

p. 5-41, ll. 15-36: A recapitulation of time averages would be helpful.
Section 5.2.2.2.: There should be some discussion of what is considered adverse or clinically relevant as was presented in the previous section.

p. 5-47, Figure 5-3: the dose level should be included here.

p. 5-57, ll. 1-3: This finding is of concern and raises a red flag.
l. 10: delete the word "strong"
ll. 26-28: given the weak associations between personal and ambient monitors this result is not surprising.

p. 5-61, Table 5-10: Where are the results for these studies?

p. 5-59, Table 5-12: Spita-Cohem et al. The odds ratios are not statistically significant, but it is reported that the personal EC measurements were associated; was this association statistically significant?

p. 5-70, Table 5-12: Delfino et al (2003): Were the co-pollutant results statistically significant?

p. 5-75: Several of the results represented here are not statistically significant. This should be explicitly indicated.

p. 5-78, l. 4: "consistency" alone is misleading. "Some consistency" would be better.

p. 5-79, Table 5-13: Present results.

p. 5-87, l. 12: "positive" but not statistically significant.
ll. 19-22: but these are not the correct co-pollutants

p. 5-92, l 11: Were they statistically significant?
Section on Seasonal Differences; Note that the correlation between personal and ambient measures also differ by season.

P. 5-93, ll. 24-34: It should be noted that measurement error will also affect the shape of the dose-response curve.

p. 5-99, ll. 9-12: See above comment.

p. 5-103, Experimental Studies: This section should also discuss alternative measures/definitions of adversity.

p. 5-108: More attention should be given to studies with exposures near the current NNAQS.
p. 5-110, Table 5-19: Were any measures of personal exposure available for any of these studies?

p. 5-14, Table 5-20: Delfino et al (20206) are NO_2 results in models with co-pollutants statistically significant?

p. 5-123, ll. 24: Here is where there should be some discussion of adversity or clinical relevance.

pp. 5-164-165: I would like to see results presented when they were considered with co-pollutants as well.

p. 5-239, line 6; delete “several”

p. 5-250: I would like to see the argument give emphasis to those studies that considered relevant co-pollutants and to those studies that made use of personal and indoor measurements of NO_2. Among human clinical studies and panel studies, I would like to see emphasis to those with exposures near ambient levels.

p. 5-303, l. 12: a result is or is not statistically significant.

p. 5-304: ll. 14-15: given this disparity can we make any inference from the results?

p. 5-305: Table 5-5: Was there any consideration of co-pollutants?”

p. 5-313, ll. 5-15: can we say anything about the clinical significance of these results?

p. 5-234, l. 15 and l. 19: Are these results statistically significant?

p. 5-328: Table 5-58: I would like to see the argument give emphasis to those studies that considered relevant co-pollutants and to those studies that made use of personal and indoor measurements of NO_2. Among human clinical studies and panel studies, I would like to see emphasis to those with exposures near ambient levels. Statistical significance should also be indicated,

p. 5-338, Table 5-60: %Increase should indicate per 20ppb increment.

P. 5-343, Table 5-62: These are not the most relevant co-pollutants.

p. 6-7, Table 6-1: Please indicate results with co-pollutants.

p. 6-18, Figure 6-2: what is meant by soot? PM2.5, EC, BC?

p. 6-46, Table 6-4: Please indicate results with co-pollutants.
p. 6-155, l. 9: statistically significant?

p. 6-173, Table 6-18: what exposure metric is the correct one? It would be useful here to see how NO_2 levels have changed since 1975. See my comments for Chapter 2.

p. 6-192. Table 6-20: I am particularly concerned by this table. Given the long latency for cancer, the reported NO_2 or NOx concentrations are probably irrelevant. Long term trends of NO_2 are needed. See my comments for Chapter 2. In addition, ambient NO_2 is associated with known carcinogens (e.g., benzo(a)pyrene); the concentrations and associations were likely much greater 20-30 years ago. Given the lack of information mentioned above, I am troubled by the "suggestive" label.