DATE

EPA-CASAC-17-XXX

Administrator E. Scott Pruitt
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 Sulfur Oxides – Health Criteria (Second External Review Draft – December 2016)*

Dear Administrator Pruitt:

The Clean Air Scientific Advisory Committee (CASAC) Sulfur Oxides Panel met on March 20-21, 2017, to peer review the EPA's *Integrated Science Assessment for Sulfur Oxides – Health Criteria (Second External Review Draft – December 2016)*, hereafter referred to as the Second Draft ISA. The CASAC's consensus responses to the agency's charge questions and the individual review comments from members of the CASAC Sulfur Oxides Panel are enclosed.

Overall, the Second Draft ISA is an improved document and is responsive to the CASAC's comments (EPA-CASAC-16-002, April 15, 2016) 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 agency's review of the Sulfur Oxides Primary (Health-based) National Ambient Air Quality Standards (NAAQS).

The CASAC finds the revised Executive Summary and Integrative Synthesis chapter to be improved. The material and format appropriately highlights and summarizes the important information provided in the subsequent chapters. A few suggestions for further clarification of the language and for additional items that could be included and highlighted in these sections are provided in the consensus responses.

The revised chapter on atmospheric chemistry and ambient air concentrations of sulfur dioxide resolves many of the inconsistencies that were found in the First Draft ISA. To improve the chapter, emission trends from recent years should be added. It is also important to highlight the contributions of emissions

from smelters and integrated iron and steel mills as these may explain some high values in the data shown. The importance of pollution sources and formation of other sulfur compounds such as inorganic and organic particulate S(IV) and organic S(VI) species should be discussed. Additional data analysis illustrating spatial and temporal variations of SO_2 concentrations would be helpful. The consensus responses contain several suggestions on ways to improve the way the data is presented and the calculation and display of the peak-to-mean ratio (PMR) data. The effects of atmospheric stability (e.g., time of day), wind speed, source type (e.g., stack height), distance from sources, and site locations on PMR should be described. These parameters should be incorporated into the EPA's empirical formula.

It should be acknowledged that AERMOD can be modified to calculate 5-minute average SO_2 concentrations. The EPA is encouraged to require local and state agencies to routinely report 5-minute SO_2 measurements, to evaluate model performance in the future. As biases in model performance have impacts on health assessment, model results should be compared to available observations and biases should be documented.

The revised chapter on exposure to ambient SO_2 is better organized and articulated than in the First Draft ISA. The new material on exposure considerations specific to sulfur oxides is helpful. Although much has improved, the chapter would benefit from additional refinement to improve clarity, language consistency, organization, and readability. The EPA is encouraged to leverage related work from recent ISAs for other criteria pollutants, particularly those more recent than the 2008 Sulfur Oxides (SO_X) ISA. By referencing and/or bringing forward previously discussed material, this document can build on the success of those previous ISAs. The consensus responses contain several specific suggestions on how this chapter can be further improved.

The modeling section (Section 3.3.2) is still difficult to understand in terms of the goals of this chapter. The chapter should clearly answer the following two questions: 1) what are the different approaches to exposure modeling? and 2) how does the selection and application of a particular exposure modeling approach affect the analysis and conclusions to be drawn from an epidemiologic study? Furthermore, exposure modeling is used in the risk and exposure assessment (REA). Thus, the chapter should also address aspects of exposure science that are relevant to application of exposure models in the REA. Cross-references to other chapters should also be added.

Overall, the revised ISA adequately characterizes the respiratory effects observed in controlled human exposure and epidemiologic studies. The descriptions of the respiratory tract, minute ventilation, and respiratory physiology associated with exercise and upper airway obstruction are well done and straightforward. The chapter is excellent in outlining the factors that impact uptake and dosimetry of SO_X and provides a thorough review of results from controlled exposure studies of adult human volunteers on the effect of SO_X on airway function, resistance, and response to allergens. The ISA would benefit from a more comprehensive conceptual description of how transport processes transform "inhaled dose" at the airway opening into "uptake" into a local target tissue.

The chapter on integrative health effects of exposure to sulfur oxides is impressive, summarizing a large and complex literature in a generally clear and efficient manner. The revised chapter addresses the

previous CASAC concerns regarding the causal determinations of the eight classes of health outcomes. The chapter now effectively presents the evidence for a causal relationship between respiratory effects and short term SO_2 exposure, based on evidence of exacerbation of asthma in both observational and experimental studies. The key role of experimental studies, which provides clear evidence of an effect of SO_2, is well described, and the crosswalk between the animal and human evidence with regard to lags and levels of exposure, the dosimetry and mode of action is compelling. The CASAC concurs with the determination that the evidence for the relationship between long-term SO_2 exposure and respiratory effects is now "suggestive of, but not sufficient to infer, a causal relationship." The evidence includes two new studies of asthma incidence in children and several experiments in rodents. A few suggestions for further improvements of the chapter are provided in the consensus responses.

The chapter on populations and lifestages potentially at increased risk for health effects related to sulfur dioxide exposure is important to the REA and Policy Assessment. It is also an important information resource for environmental management policies, public health organizations, and the public. The introduction for this chapter needs to provide an expanded and clear discussion of the objectives of this chapter and its implications for subsequent documents. A more comprehensive discussion of all the factors that are associated with increased risk would also improve the chapter. Details are provided in the consensus responses.

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

Sincerely,

Ana V. Diez Roux, 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
Sulfur Oxides Panel

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

CASAC MEMBERS
Dr. Judith Chow, Nazir and Mary Ansari Chair in Entrepreneurialism and Science and 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. Donna Kenski, Data Analysis Director, Lake Michigan Air Directors Consortium, Rosemont, IL

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

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

Dr. John R. Balmes, Professor, Department of Medicine, Division of Occupational and Environmental Medicine, University of California, San Francisco, San Francisco, CA

Dr. James Boylan, Program Manager, Planning & Support Program, Air Protection Branch, Georgia Department of Natural Resources, Atlanta, GA

Dr. Aaron Cohen, Consulting Scientist, Health Effects Institute, Boston, MA

Dr. Alison C. Cullen, Professor, Daniel J. Evans School of Public Policy and Governance, University of Washington, Seattle, WA

Dr. Delbert Eatough, Professor of Chemistry, Department of Chemistry and Biochemistry, Brigham Young University, Provo, UT

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

Dr. William C. Griffith,* Associate Director, Department of Environmental and Occupational Health Sciences, Institute for Risk Analysis & Risk Communication, School of Public Health, University of Washington, Seattle, WA

Dr. Steven Hanna, President, Hanna Consultants, Kennebunkport, ME

Dr. Daniel Jacob,* Professor, Atmospheric Sciences, School of Engineering and Applied Sciences, Harvard University, Cambridge, MA

Dr. Farla Kaufman, Epidemiologist, Office of Environmental Health Hazard Assessment, Reproductive and Cancer Hazards Assessment Section, California EPA, Sacramento, CA

Dr. David Peden, Distinguished Professor of Pediatrics, Medicine & Microbiology/Immunology, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

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

Dr. Frank Speizer, Edward Kass Distinguished Professor of Medicine, Channing Division of Network Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA

Dr. James Ultman, Professor, Chemical Engineering, Bioengineering Program, Pennsylvania State University, University Park, PA

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

* Did not participate in review

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

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

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

Dr. Ivan J. Fernandez, Distinguished Maine Professor, School of Forest Resources and Climate Change Institute, University of Maine, Orono, ME

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

Dr. Donna Kenski, Data Analysis Director, Lake Michigan Air Directors Consortium, Rosemont, IL

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. 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 the EPA's *Integrated Science Assessment for Sulfur Oxides – Health Criteria (Second External Review Draft – December 2016)*

Charge #1 – Executive Summary and Chapter 1

Please comment on the extent to which revisions to the Executive Summary and Chapter 1 have reduced redundancy and made the Executive Summary more accessible to a nontechnical audience.

In the CASAC's review (EPA-CASAC-16-002, April 15, 2016) of the First Draft ISA, the following comments and suggestions for the Executive Summary (ES) and Chapter 1 were made:

- Consider revising the language for a broader, non-technical audience and eliminating technical jargon as much as possible;
- Clearly state in the Executive Summary (ES) and Chapter 1 that the controlled human exposure studies are the principal rationale behind the 2010 1-hour SO_2 NAAQS and that this standard also provides protection from chronic exposure effects;
- Summarize the correlation of maximum 5-minute SO_2 concentrations with corresponding 1-hour concentrations;
- Clearly and consistently define short- and long-term exposures throughout the text of the entire ISA, including the ES;
- Elevate some of the footnotes used in the first page of the ES to the body of the text; and
- Mention ambient background concentrations of SO_2 in this summary section of the ISA.

Overall, the CASAC finds that the Second Draft ISA adequately addresses the CASAC's comments and suggestions for the ES and Chapter 1. The ES is now reasonably free from technical jargon. The EPA is still encouraged to further refine this important section of the ISA so that it is can be more understandable (readable) for a wider sector of the public. As an example, to be less ambiguous, the opening sentence of the ES could be changed to:

> "This Integrated Science Assessment (ISA) is a comprehensive evaluation and synthesis of policy-relevant science aimed at 1) characterizing exposures to <u>ambient sulfur dioxide</u> (SO_2), <u>the primary atmospheric indicator of</u> <u>gaseous sulfur oxides</u> (SO_X), and 2) the health effects associated with these exposures."

In addition, the CASAC suggests replacing the term "sulfur aerosols" with "<u>particles of sulfur oxides</u>." Use of the term "SO_X" elsewhere in the ES such as page xliii, lines 14-16, could also be clarified.

The CASAC finds that the material and format in the ES and Chapter 1 appropriately highlights and summarizes the important information provided in the subsequent chapters. Examples of integration and synthesis of the chapters is somewhat limited in this opening introduction to the ISA, but are most effectively captured in the sections on causality determinations. The causality determination for the

respiratory health effects related to short-term sulfur dioxide (SO_2) exposures has been appropriately highlighted in both the ES and Chapter 1. However, the one change in causality determination (respiratory effects related to long-term SO_2 exposure) from the 2008 Integrated Science Assessment for Sulfur Oxides – Health Criteria (USEPA, 2008; hereafter referred to as the 2008 ISA) needs to be more clearly highlighted in Table 1-1.

Now that a more robust ambient 5-minute dataset is available, the CASAC suggests including some discussion on how often 5-minute SO_2 concentrations over the 200 - 400 ppb range of concern actually occurs on an annual basis (occurrence and frequency of potentially harmful short-term, 5-minute exposures) in the national database.

Charge #2 – Atmospheric Chemistry and Ambient Concentrations of Sulfur Dioxide and other Sulfur Oxides

Please comment on the extent to which these revisions improve the characterization of sources, chemistry, and concentrations of ambient sulfur oxides and hence provide a scientific foundation for subsequent technical and policy analyses during the review of the SO_2 NAAQS.

Sources of Sulfur Dioxide

In the First Draft ISA, the CASAC found that the source categories and definitions of major sources were inconsistent. The Second Draft ISA addresses this inconsistency and the 12 major SO_2 source categories are now consistent. The emission trends (Table 2-1 and Figure 2-5) should include years 2011 to 2016 to reflect more recent emission estimates—a reduced number of source categories may be considered for clarity when plotting the emission trends. There have been major SO_2 reductions over the past few years, especially for electric generating units (EGUs). A table summarizing locations and SO_2 emission rates for metals processing subcategories (e.g., copper/lead smelters and integrated iron and steel mills) should be included to explain the 107,000 tons/year emissions in 2016. It should be noted that motor vehicle engine exhaust contributed only ~2-3% of SO_2 in the recent national emissions inventory. The data in Figure 2-11 and Figures 2-13 through 2-18 indicate that some monitors influenced by anthropogenic SO_2 emissions with a 99th percentile 5-minute maximum concentration in the range of >75 to 200 ppb or >200 to 400 ppb (e.g., the highest concentrations to which a population is currently exposed) are dominated by emissions from copper/lead smelters and integrated iron and steel mills. Tables should be added to each of these figures to identify the locations of the monitors with the highest concentrations and possible influences from nearby industrial sources should be discussed.

Atmospheric Chemistry and Fate

The importance of pollution sources and formation of other sulfur compounds such as inorganic particulate S(IV) species, organic particulate S(IV) species (e.g., bis-hydroxy dimethyl sulfone), and organic S(VI) species (e.g., alkyl sulfates) should be discussed. Past toxicological studies regarding the synergistic health effects of SO_2 and S(IV) should be examined and compared with SO_2 inhalation

response (e.g., Alarie et al., 1973; Amdur, 1971). These compounds were potential confounders or moderators of SO_2 health effects in epidemiologic studies where copper smelter or integrated iron and steel mill emissions were present. Relevant health studies should be reviewed and the potential influence of these non-sulfate compounds should be noted.

Environmental Concentrations

For concentrations below instruments' lower detection limits (LDLs), negative values represent deviations about the instrument baseline that offset positive deviations. The elimination of data below zero will bias SO_2 concentration distributions (Table 2-6) when many concentrations are below the LDLs.

The importance of and corrections to water vapor induced collisional quench in the pulsed fluorescence Federal Reference Method (FRM) needs elaboration. The cited Luke (1997) experiment was conducted for 0.5 ppmv SO_2 (well below FRM LDLs) and was deemed by the author "…highly uncertain…due to lack of extensive data."

SO_2 instruments have different time constants (e.g., 10 – 300 seconds) to reduce noise. Those with time constants on the order of 300 seconds will smear the peaks over adjacent 5-minute intervals. This should be recognized and a recommendation should be made for <5-minute instrumental smoothing while still maintaining <2 ppb LDLs.

Table 2-6 also shows that 5-minute hourly maximum values were lower than those of 1-hour daily maximum values for 2013 and 2015, inconsistent with the trend in the 99th percentile distributions. A lognormal distribution for 5-minute hourly maximum values versus 1-hour average SO_2 concentrations may be considered, in addition to the scatterplot shown in Figure 2-26, to better represent pollution concentration distributions (USEPA, 1974). The same approach can be adapted to fill in missing data for modeling and to illustrate inter-site pairwise correlations with distance (i.e., Figures 2-19 and 2-20). In addition to the six focus areas, characteristics of 5-minute exposures of concern at the national scale should be addressed.

Additional data analysis to illustrate spatial and temporal variations of SO_2 concentrations is needed. As spatial variability of pollutant concentrations are characterized by random turbulence, the meso-distance scale of atmospheric turbulence (e.g., tens of kilometers) should be considered. (Please see Dr. Steven Hanna's individual comments for additional details). Diurnal variations of SO_2 concentrations near large source regions should be shown to demonstrate the importance of temporal variations.

Peak-to-mean ratios (PMRs) have been used to evaluate the distribution of 5-minute hourly maximum SO_2 concentrations corresponding to a given hourly value. As PMR increases with longer averaging times, the denominators used to calculate PMRs need to be clarified. The effects of atmospheric stability (e.g., time of day), wind speed, source type (e.g., stack height), distance from sources, and site locations on PMR should be described. These parameters should be incorporated into the EPA's empirical formula to demonstrate their variations.

Atmospheric Modeling

Science algorithms in Gaussian models apply to several averaging times, but the ISA focuses on hourly AERMOD modeling. The ISA should acknowledge that AERMOD can be modified to calculate 5-minute average SO_2 concentrations. The EPA is encouraged to require local and state agencies to routinely report every 5-minute measurement in the future. As biases in model performance have impacts on health assessment, model results should be compared to available observations and biases should be documented.

Charge #3 – Exposure to Ambient Sulfur Dioxide

Please comment on these revisions insofar as the chapter supports the evaluation of the strength of inference and potential for copollutant confounding in epidemiologic studies in Chapter 5.

The revised chapter on exposure assessment is better organized and articulated than in the First Draft ISA. The new material on exposure considerations specific to SO_X is helpful. Although much has improved, the chapter still needs considerable refinement, including improved clarity, organization, and readability. The EPA is encouraged to leverage related work from other recent ISAs for other criteria pollutants, particularly those more recent than the 2008 SO_X ISA. By referencing and/or bringing forward previously discussed material, this document can build on the success of those previous chapters.

There is a lack of clarity in the definition and usage of multiple terms, and some terms are defined incorrectly or used inconsistently. A few examples are provided here. (Please see the extensive individual comments given by Drs. Cullen, Frey and Sheppard for additional details.) For instance, in Section 3.2.1 the terms “exposure metric,” “surrogate,” “differential misclassification,” “nondifferential misclassification,” “exposure error,” “bias,” and “precision” should be discussed more accurately. It is not always clear when the focus of the text is on exposures themselves (e.g. exposures estimated from models or obtained directly from measurements) or on the target parameter of interest in an epidemiologic study. Please clarify whether the phrase “surrogate parameter of interest” is the same as “surrogate target parameter of interest,” and please articulate the purpose of adding the term “surrogate” to the idea of a (target) parameter of interest. Sometimes the same term is used for two distinct concepts, such as uncertainty “in” or “about” the exposure metric. In this example, language about uncertainty in the estimate itself, and language about uncertainty relative to which metric to use, are conflated. Distinguishing bias and precision is also recommended as these terms are conflated in places. Please clarify that exposure assessment includes both exposure design and measurement as well as exposure estimation. Replacing the term “central site monitor” with the term “fixed site monitor” is suggested.

Equation 3-1 has mixed indexing for the integral. It is necessary to clarify whether integration is across *j* (the microenvironments) or *t*, in which case, the equation is estimating E_j (rather than E_T).

The modeling section (Section 3.3.2) is still difficult to understand in terms of the goals of this chapter. The chapter should clearly answer the following two questions: 1) what are the different approaches to exposure modeling? and 2) how does the selection and application of a particular exposure modeling approach affect the analysis and conclusions to be drawn from an epidemiologic study? (See Dr. Sheppard's individual comments for detailed comments on this point.) Furthermore, exposure modeling is used in the risk and exposure assessment (REA). Thus, the chapter should also address aspects of exposure science that are relevant to application of exposure models in the REA.

The CASAC had previously suggested adding new tables to this chapter. Table 3-1 is a good addition that focuses on the epidemiologic perspective. However, it needs some refinement: more accuracy and specificity is needed in some places, while extraneous details should be removed in other places. (Please see Dr. Frey's individual comments for specific suggestions.) It would be helpful to add text to Section 3.3.3 that provides a road map to understanding each column of Table 3-1. For instance, the exposure assignment method column lumps into one category the source of data and the method for obtaining an estimated exposure. The column addressing errors and uncertainties should clearly articulate whether it is describing the impact on the exposure estimate itself or the impact of using this estimated exposure on epidemiologic inference; also this column should address sources of bias and imprecision separately rather than conflating them. The CASAC suggests that only exposure assignment methods that can be referenced in at least one study reviewed in Chapter 5 of the current ISA or reviewed in the 2008 ISA be included in this table. The document would benefit from two additional tables. Adding a table to complement Table 3-1, from the exposure perspective, that compares the modeling approaches based on their characteristics is suggested. (Please see the individual comments from Dr. Frey for details.) Another suggestion is to consider a glossary or a tabular presentation with all the terms defined from Section 3.2.1. For instance, the World Health Organization has documents that include this type of glossary (see e.g., WHO, 2004); these could be used as examples.

Section 3.4 should introduce the concept of study design before addressing details about factors contributing to error in exposure. Details about study designs that are not relevant to material in Chapter 5 of the current ISA or to the 2008 ISA do not need to be covered in any depth (e.g. panel studies). The agency is encouraged to continue to focus on and improve two sections: approaches to exposure modeling (Section 3.3.2), and the impact of exposure estimates on conclusions from epidemiologic studies (Section 3.4; specifically, Section 3.4.4). An additional section should be added to explain key similarities and differences in exposure modeling for the REA versus the impact of exposure estimates on epidemiology.

The treatment of air exchange rate, penetration factor, and deposition rate should be more logically structured and organized. The discussion of factors affecting air quality inside vehicles requires acknowledgment that factors can be interactive depending on cabin ventilation (e.g., windows open). The discussion of use of GPS for activity tracking requires a more up-to-date treatment.

The chapter does not touch upon microenvironmental monitoring as a method for measurement of exposure concentrations. Either such material could be added or, if such methods are not typically applied for SO_X, then a brief statement to this effect would be helpful.

The agency is encouraged to add cross-referencing to other chapters in the ISA and to also anticipate references to this chapter in future REA and PA documents, as appropriate. For instance, this chapter should refer to Chapter 2 for information about sources of SO_2. All references to Chapter 3 that appear in Chapter 2 should also be reviewed and ensure that the reverse pointers to Chapter 2 also appear in Chapter 3. For example, the bottom of p. 2-29 refers to Section 3.2.1, but the reverse reference to Chapter 2 is absent from Chapter 3. As a further need for cross-referencing, the points made in Tables 5-21, 5-24, and 5-41 about exposure measurement error and co-pollutant effects should be clearly identifiable in Chapter 3. Thus the agency should ensure Chapter 3 text refers to these Chapter 5 tables and related text. The agency is encouraged to ensure that ideas expected to be included in the REA or PA are covered reasonably in the ISA so that these future documents can easily reference appropriate sections of the ISA. As an example, the scientific basis of methods for imputing missing 5-minute ambient concentrations should be addressed in the ISA.

Charge #4 – Health Effects of Short-Term and Long-Term SO_2 Exposure

Please comment on the extent to which our characterization of the evidence and rationale for these causal determinations is consistent with the EPA's causal framework.

Please comment on the adequacy of the characterization of respiratory effects observed in controlled human exposure and epidemiologic studies, particularly in different populations and lifestages.

Overall, Chapters 4 and 5 adequately characterize the respiratory effects observed in controlled human exposure and epidemiologic studies. Chapter 4 is well written and the description of the respiratory tract, minute ventilation and respiratory physiology associated with exercise and upper airway obstruction made reviewing the dosimetry of SO_X much more straightforward. Chapter 4 is excellent in outlining the factors that impact uptake and dosimetry of SO_X, including the effect of respiratory rate and upper airway physiology and obstruction in delivering this gas to the lower airway and the chemistry and absorption of this gas into the lower airway, as well as the metabolism of this gas once taken up systemically. The ability of the gas to activate sensory nerves (and impact lung physiology) and induce airway injury at high doses is well covered. This chapter also provides a thorough review of results from controlled exposure studies of adult human volunteers on the effect of SO_X on airway function, resistance, and response to allergen in allergic and allergic asthmatic volunteers. However, diffusion-reaction processes that contribute to local uptake between the gas phase and the epithelial lining fluid (ELF) are not adequately discussed.

Some of the key findings from Chapter 4 include:

- SO_X exerts effects as rapidly as 5 minutes after initiation of an exposure;
- SO_X effects include immediate effects on lung function that suggest a sensorineural impact on respiratory mechanics, which is pronounced in asthmatic volunteers;
- Exposure to SO_X enhances response to allergen challenge in allergic individuals, either alone or in conjunction with co-pollutants;

- There is an inflammatory effect to SO_X, and inflammatory response to allergen (in allergic volunteers) is enhanced following SO_X exposure;
- The ability of beta adrenergic agents, anticholinergic agents, leukotriene modifiers and inhaled corticosteroids on response to SO_X suggest both sensorineural as well as inflammatory mechanisms on adverse health outcomes to this gas, though in asthmatics, none of these agents completely ablates these responses;
- Response to SO_X varies between individuals- there are responders and non-responders even within asthmatic groups, suggesting mechanisms independent of asthma/atopy that determine response to this gas (this point needs to be more explicitly stated in the document);
- To date, there are minimal findings for extrapulmonary effects of exposure to SO_X;
- Body mass index does modify response to inhaled pollutants and is likely an important determinant of response to SO_X.

To improve the document, specific recommendations for revision of Chapter 4 are outlined below.

<u>Section 4.2.2-Absorption</u>

The revised ISA should include a more comprehensive conceptual description of how transport processes transform "inhaled dose" at the airway opening into "uptake" into a local target tissue. This transformation involves longitudinal convection and diffusion processes in the respired gas phase as well as lateral diffusion and reaction processes in the underlying ELF. In a simple model, local uptake is proportional to the difference in local pollutant concentration between the gas and tissue phases. The proportionality constant is an overall mass transfer coefficient that depends on gas-phase mass transfer coefficient, physical solubility, liquid-phase molecular diffusion coefficient, liquid phase reaction rate coefficient and ELF layer thickness (e.g., Hu et al., 1992).

Estimation of SO_2 absorption rates in children and young adults (pg 4-9, lines 3-5) make the implicit assumption that the transport resistance to lateral SO_2 is controlled by the gas phase. The assumption should be explicitly stated and justified. A more complete description of the uptake process, which recognizes the roles of both gas-phase and liquid-phase transport processes (see above paragraph), would help in this regard.

This section (inadvertently) gives the impression that anatomical and functional factors are the only differences between children and adults. It should also be recognized that ontological factors like airway development can be important considerations (e.g., Foos et al., 2008).

<u>Section 4.3-Mode of Action of Inhaled Sulfur Dioxide</u>

On line 10 of page 4-18, the statement "SO_2 is a highly reactive antioxidant gas" should likely read "SO_2 is a highly reactive gas" or "SO_2 is a highly reactive oxidant gas." In addition to oxidative stress, the impact of H+ (acid state) on airway physiology should be acknowledged.

Although there is a robust description of experimental studies of the effect of SO_2 in human volunteers and asthmatics in particular, it should be more explicitly stated that some persons are regularly responsive to SO_2 and others are non-responsive, independent of other co-morbidities, suggesting specific biological determinants for SO_2 response.

Chapter 5

Chapter 5 reviews and summarizes the evidence from epidemiologic, controlled human exposure and animal studies regarding the adverse health effects of exposure to SO_X. Chapter 5 is impressive, summarizing in a generally clear and efficient manner, a large and complex literature and its evolution since the 2008 ISA. The links to Chapter 3 (exposure) and Chapter 4 (dosimetry) are very helpful and serve well to buttress arguments made concerning the health evidence. The revised chapter largely addresses the previous CASAC concerns regarding the causal determinations of the eight classes of health outcomes.

The chapter now effectively presents the evidence for a causal relationship between respiratory effects and short term SO_2 exposure, based on evidence of exacerbation of asthma in both observational and experimental studies. The key role of experimental studies, which provides clear evidence of an effect of SO_2, is well described, and the crosswalk between the animal and human evidence with regard to lags and levels of exposure, the dosimetry and mode of action (Chapter 4) is compelling (see above).

With regard to long-term SO_2 exposure and respiratory effects, the determination that the evidence is now "suggestive of, but not sufficient to infer, a causal relationship" rather than "inadequate to infer the presence or absence of a causal relationship" rests on two new studies of asthma incidence in children and several experiments in rodents. The experimental studies are key because they serve to mitigate, to some extent, concerns with attributing the observed effects to SO_2, as opposed to, for example, sulfate particles, given that neither epidemiologic study provided information on co-exposures.

The revised cardiovascular disease section provides a more critical and balanced assessment of the available evidence, including a more critical assessment of the potential for copollutant confounding. Despite the accrual of many new studies since the 2008 ISA, the evidence remains "inadequate to infer the presence or absence of a causal relationship" between cardiovascular effects and both long-term and short-term SO_2 exposure. The contrast with the evidence regarding the relationship between short-term SO_2 exposure and respiratory effects, is notable, especially the relative lack of support from experimental studies.

The CASAC concurs that the evidence for reproductive and developmental effects is "inadequate to infer the presence or absence of a causal relationship." Although there were some epidemiologic studies that observed reproductive effects, copollutant confounding was not addressed (these studies also observed effects of PM, CO and NO_2 on adverse reproductive outcomes). Moreover, there was a lack of experimental evidence to support an association.

The revised chapter provides stronger support for the determination that the evidence is "suggestive of, but not sufficient to infer, a causal relationship" between short-term SO_2 exposure and total mortality. The chapter more clearly documents that, despite the accrual of more studies, the shortcomings of the literature identified in the 2008 ISA have largely not been addressed. Specifically, the potential for copollutant confounding remains largely unresolved. To the extent that copollutant confounding has been assessed, neither $PM_{2.5}$ nor sulfate PM has been included, which is a particular issue in studies conducted in China, where coal burning is the major source of pollution.

The CASAC concurs that the evidence is "inadequate to infer the presence or absence of a causal relationship" between mortality and long-term SO_2 exposure and between cancer and long-term SO_2 exposure.

Some of the key findings from Chapter 5 include:

- Controlled human exposure and observational studies indicate that short-term exposure to SO_2 causes increased severity of disease in asthmatics;
- Children with asthma are at particular risk of exacerbation of their disease due to short-term exposure to SO_2;
- These observations are supported by experiments in animals and by mechanistic evidence;
- Individuals with asthma display a wide range of responsiveness to SO_2;
- The evidence of a causal relationship between SO_2 exposure and cardiovascular, and reproductive and developmental effects is considerably weaker due to copollutant confounding and lack of evidence from experimental and mechanistic studies.

To improve the document, suggested revisions for Chapter 5 include:

- Consistent language should be used throughout the document to describe uncertainty in findings, statistical significance vs. biological impact, and impact on health effects;
- Confidence intervals give a sense of the precision of an estimate, but are too often inappropriately used as tests of significance, i.e., "imprecision" is used as a code word for a confidence interval includes the null value. This practice should be avoided and the term "imprecise" should be correctly used to refer to the width of the interval;
- Error in exposure estimates should be clearly acknowledged;
- Most controlled exposure studies employ exercise and effects are rarely seen in the lower airway without exercise. For this reason, exercise during exposure to ambient SO_2 is an important consideration in determining the risk for SO_2 induced adverse health outcomes and needs to be clearly discussed.
- Asthmatics are a key risk group. It should be noted that most of the epidemiologic studies and controlled exposure studies have focused on children with asthma and young adults with asthma. Most of these people have allergic asthma. The effects of exposure to SO_2 in severely affected asthmatics and persons with different phenotypes (variants) of asthma or other airway diseases (e.g. COPD) are less well studied;

- Although the prevalence of asthma is approximately 7-8% in the general population; this varies with respect to sex, age and ethnicity (e.g. approximately 15% in non-Hispanic African American children), thus, certain segments of the population are at increased risk;
- Although this chapter clearly fulfills the charge, it is an exhaustive chapter with many tables that might be improved by focusing on the summary statements and using the tables as supplemental or referenced information, and further integrating the findings.

Charge #5 – Populations and Lifestages Potentially at Increased Risk for Health Effects Related to Sulfur Dioxide Exposure

Please comment on the adequacy of these revisions to clarify the characterization of the evidence for increased risk of SO_2-induced health effects in different populations and lifestages.

Chapter 6 is an important chapter that provides a link between the ISA, the REA, and the Policy Assessment. It is also an important information resource for environmental management policies, public health organizations, and the public. The challenge is that the material in this chapter is drawn from other parts of the ISA, and there needs to be a way to integrate this material to provide value without being repetitious. Currently this chapter makes a good initial attempt to do this, but it needs to be revised and strengthened. In particular, the introduction for this chapter needs to provide an expanded and more articulate discussion of all of the objectives of this chapter and how its content will be used.

This chapter needs to more clearly articulate the factors that are associated with increased risk. Human clinical studies show that a portion of asthmatics, but not all, appear to be sensitive to low levels of SO_2 exposure. It is unclear what factors lead to this difference in response; it could be asthma phenotype or some genetic factor. Other factors that can influence asthmatic response are tied to greater ventilation rates, such as increased exercise, obesity, and age, particularly children. Asthma prevalence is also greater among certain subpopulations, such as black children; hence this subgroup may be at greater risk. Other factors that have been shown to increase asthmatic response according to chamber studies include weather, particularly cold weather and absence of asthma medication. The current chapter needs to identify all of these factors more comprehensively. In addition, cumulative risk factors could influence responses, such as exposure to more than one pollutant or to an allergen, smoking, and other compromised health issues.

References

Alarie, Y., Busey, W. M., Krumm, A. A., and Ulrich, C. E. (1973). Long-Term continuous exposure to sulfuric acid mist in cynomolgus monkeys and guinea pigs. *Archives of Environmental Health: An International Journal*, 27(1):16-24.

Amdur, M. O. (1971). Aerosols formed by oxidation of sulfur dioxide. *Archives of Environmental Health: An International Journal*, 23(6):459-468.

Foos, B., Marty, M., Schwartz, J., Bennett, W., Moya, J., Jarabek, A. M., and Salmon, A. G. (2007). Focusing on children's inhalation dosimetry and health effects for risk assessment: An introduction. *Journal of Toxicology and Environmental Health, Part A*, 71(3):149-165.

Hu, S.-C., Ben-Jebria, A., and Ultman, J. S. (1992). Simulation of ozone uptake distribution in the human airways by orthogonal collocation on finite elements. *Computers and Biomedical Research*, 25(3):264-278.

Luke, W. T. (1997). Evaluation of a commercial pulsed fluorescence detector for the measurement of low-level SO2 concentrations during the Gas-Phase sulfur intercomparison experiment. *Journal of Geophysical Research: Atmospheres*, 102(D13):16255-16265.

U.S. EPA (1974). *Proceedings of the Symposium on Statistical Aspects of Air Quality Data.* Meteorology Laboratory, National Environmental Research Center, Office of Research and Development, U.S. Environmental Protection Agency, Research Triangle Park, N.C., EPA-650/4-74-038, October 1974.

U.S. EPA (2008). *Integrated Science Assessment (ISA) for Sulfur Oxides – Health Criteria* (Final Report, Sep. 2008). U.S. Environmental Protection Agency, Washington, DC, EPA/600/R-08/047F, 2008.

World Health Organization (WHO) (2004). *IPCS Risk Assessment Terminology*. International Programme on Chemical Safety, World Health Organization, Geneva, Switzerland, 2004. Available at: http://www.inchem.org/documents/harmproj/harmproj/harmproj1.pdf.

Appendix A

Individual Comments by CASAC Sulfur Oxide Panel Members on the EPA's *Integrated Science Assessment for Sulfur Oxides – Health Criteria (Second External Review Draft – December 2016)*

Mr. George Allen **A-2**
Dr. John Balmes **A-4**
Dr. James Boylan **A-7**
Dr. Judith Chow **A-9**
Dr. Aaron Cohen **A-12**
Dr. Alison C. Cullen **A-16**
Dr. Delbert Eatough **A-20**
Dr. H. Christopher Frey **A-28**
Dr. Steven Hanna **A-39**
Dr. Jack Harkema **A-43**
Dr. Farla Kaufman **A-45**
Dr. Donna Kenski **A-48**
Dr. Richard Schlesinger **A-50**
Dr. Elizabeth A. (Lianne) Sheppard **A-51**
Dr. Frank Speizer **A-59**
Dr. James Ultman **A-67**
Dr. Ronald Wyzga **A-69**

Mr. George Allen

Charge #1

Please comment on the extent to which revisions to the Executive Summary and Chapter 1 have reduced redundancy and made the Executive Summary more accessible to a nontechnical audience.

Overall the revisions to these sections are responsive to the comments in the CASAC review of the first draft of the SO_X ISA. The Executive Summary is now reasonably free from technical jargon. Table ES-1 is a concise summary of health effect causal determinations; it reflects the requested changes for several categories from suggestive to inadequate, leaving only long-term respiratory effects as a change from the last NAAQS review (from inadequate to suggestive). Table 1-1, the more detailed version of this table in the Integrated Synthesis, is well done.

I still have some concerns with how the term SOx is used in the Executive Summary and Chapter 1. As used in this ISA, SOx only means gaseous SOx (SO2 and SO3, not sulfur oxides in the particle phase). However, the first paragraph of the Executive Summary uses “sulfur oxides” to describe both gas phase SOx (SO2 health effects, first sentence) and deposition of SOx (welfare effects, last sentence), which is of course sulfate. For a general audience, this could be confusing. Chapter 1 Section 1.1 first paragraph and elsewhere has similar usage.

The Chapter 1 discussion of observed ambient SO2 concentrations from the national monitoring network (Section 1.4.1, pg. 1-8, lines 1-15) is similar (the same level of minimal detail) as the corresponding paragraph in the ES. Some discussion here of how common 5-minute SO2 concentrations over the 200 to 300 ppb range of concern occur would be helpful here, now that a more robust ambient 5-minute dataset is available; this discussion could parse out the Hawaii volcano SO2 data separately since that is an unusual and uncontrollable “natural” case.

Comments on Chapter 2 - Atmospheric Chemistry and Ambient Concentrations

Section 2.4, Measurement Methods. Section 2.4.1.2 is a good discussion of the various potential interferences in the pulsed fluorescence (UVF) FRM. This section mentions water vapor quenching (page 2-25) but not how that is avoided; it only mentions permeation dryers in the context of condensation on sample inlet lines and particle filters, something that is not common practice. Current UVF FRM analyzers are reasonably free of water vapor quenching artifacts.

Measurement of peak 5-minute ambient concentrations is of interest here, but there is no mention of instrument response times in this section. Despite the FRM specifications for minimum "Rise and Fall" times of 2 minutes (Table 2-3), current monitoring regulations allow the use of a wide range of instrument "averaging times," usually anywhere between 10 to 300 seconds (5 minutes), since they assume the parameter of interest is a 1-hour average. This is noise (low-pass) filtering at the instrument (not related to any averaging done in data acquisition systems), and allowable settings are specified in the List of Designated Reference and Equivalent Methods at https://www3.epa.gov/ttn/amtic/criteria.html. An instrument averaging time of 5 minutes could result in under-reported 5-minute peak values by as much as a factor of 2. Combining this with the practice of block 5-minute averages instead of running averages based on 1-minute data could result in further under-reporting of the maximum 5-minute hourly concentration of interest for health effects. As peak to mean ratios increase, the potential for under-reporting true 5-minute concentrations with current practices also increases. Some discussion of this potential for under-reporting 5-minute peak concentrations would be appropriate in this chapter.

Table 2-6 (page 2-35): The SO2 concentrations reported in the "max" column are very high. It would be helpful if a brief note could describe the two sites driving these values. One is (as expected) a Hawaii volcano site. The other with a 5-minute max of over 4 ppm is a sulfuric acid plant in Pocatello, ID.

Section 2.5.2.2, Urban Spatial Variability, describes ambient data from six focus sites. The maps and tables do a good job of describing SO2 sources and monitors in these focus areas. It may be worth noting that these are not the focus areas that will be used in the REA – the criteria for selecting those are very different.

Section 2.5.4, Relationships between Hourly Mean and Peak Concentrations.
Figure 2-26 (Scatterplot of 5-minute hourly max versus 1-h avg) is visually misleading since there are ~ 10 million data pairs, and most of them are obscured. This plot only conveys useful information when 5-minute peak concentrations are higher than ~ 300 ppb. It would be helpful if some other approach to visually summarize these data was used. The background material on peak to mean ratios (PMR) is useful and appropriate. This section presents summary information for the six focus sites, but doesn't address characteristics of 5-minute exposures of concern at the national scale in sufficient detail. Appendix B of the REA planning document has useful summaries of this that could be useful in this section.

Dr. John Balmes

Charge # 1 - Executive Summary and Chapter 1

Please comment on the extent to which revisions to the Executive Summary and Chapter 1 have reduced redundancy and made the Executive Summary more accessible to a nontechnical audience.

I think the revised Chapter 1 is much improved regarding reduced redundancy and greater clarity for a non-technical audience.

I must say I wish I could say the same for Chapter 5. To me, Chapter 5 reads like an old Criteria Document chapter with much redundancy and excessive detail about some studies that makes the chapter harder to read and is unnecessary. I would like crisp summaries of the relevant literature in appropriate categories with thoughtful integration at the end of each category. The integration tends to be there and for that I applaud the authors, but is it really necessary to take this many pages to cover the SOx health effects literature?

p. 1-19, lines 14-16 “The limited and inconsistent evidence for these 14 nonasthma-related respiratory effects does not contribute heavily to the causal determination.” I don’t think ‘heavily’ is appropriate here. Does this limited and inconsistent evidence contribute at all to the causal determination? I would say not.

Charge # 4 - Health Effects of Short-Term and Long-Term S02 Exposure

Please comment on the extent to which our characterization of the evidence and rationale for these causal determinations is consistent with the EPA's causal framework.

In general, I approve of the characterization of the evidence and stated rationale for causal determinations that are made. I think they follow the guidance of EPA’s causal framework.

One specific problem I have with the presentation of the evidence for respiratory effects among healthy individuals is the inclusion of animal models of allergic sensitization (pp. 5-129 and 5-130). While the animals started off “naïve” and healthy, they were made to be models of human allergy. These animal models of allergic disease use experimental protocols characterized by conditions that are very different from those by which humans acquire allergic sensitization. I don’t have a problem with noting in an integrative discussion that SO2 exposure has been shown to enhance allergic sensitization in animal models, but don’t think a whole section under respiratory effects in healthy individuals is appropriate.

Please comment on the adequacy of the characterization of respiratory effects observed in controlled human exposure and epidemiologic studies, particularly in different populations and life stages.

In general, I think the respiratory effects of SO2 are appropriately characterized in the discussion of controlled human exposure and epidemiological studies.

Specific Comments

5-4, line 9 I don't know what is meant by "discipline" here.

5-39, line 19 Since the Balmes et al., 1987 study is mentioned here in the text, I think it should also be included in Table 5-7.

5-58, line 30 Should provide the actual value (75 ppb 1-hour max) here.

5-72, line 4 Should be "infiltration *of airway epithelium by* inflammatory cells…"

5-72, lines 7-8 It might be helpful to add that inflammatory cells, especially neutrophils, that infiltrate the airways due to the chemoattractant cytokine signaling generate oxidative stress.

5-72, line 25 Should spell out nasal lavage fluid here.

5-79, line 4 Should be "toward Th2 in rats…"

5-79, line 29 The evidence that SO_2 causes eosinophilic inflammation is pretty limited.

5-82, line 19 The "2-h avg metric *used in this study* is more comparable…"

5-83, line 28 "COPD is a lung disease by *destruction of alveolar tissue, airway remodeling,* and airflow limitation."

5-84, line 19 Instead of "recent results" I would say "results of the more recent study."

5-91, lines 5-9 Citations should be provided for the specific data presented, i.e., a correlation coefficient and 95% CIs.

5-120, line 6 Rather than state that "The association was imprecise…" it would be better to state "There was no significant association."

5-128, line 24 Should be "infiltration *of airway epithelium by* inflammatory cells…"

5-129-130 Why are studies of sensitized animals included in a section on normal, healthy humans?

5-132, line 5 Should be "In contrast, *Li et al.* found…"

5-134, lines 16-19 This sentence needs clarification. It suggests that there is seasonal variation, but also states that the all-year and winter results are similar to those of the summer.

5-144, line 4 Spell out IDV

5-157, line 28 Spell out TCHS

5-196, lines 2-6 Is the all-year effect really 92.9%? such a result is not similar to the summer effect of 9.4%.

5-201, lines 27-30 There is a lack of clarity across these two sentences: "**Both studies found no change in heart rate** during or immediately **following** similar exposure conditions. Tunnicliffe et al. (2001**) did not obtain electrocardiographic (ECG) measures following exposure** and thus may have been unable to capture the decrease in heart rate reported by Routledge et al. (2006)."

Dr. James Boylan

Chapter 2 - Atmospheric Chemistry and Ambient Concentrations of Sulfur Dioxide and other Sulfur Oxides

Please comment on the extent to which these revisions improve the characterization of sources, chemistry, and concentrations of ambient sulfur oxides and hence provide a scientific foundation for subsequent technical and policy analyses during the review of the SO_2 NAAQS.

The revisions to Chapter 2 have significantly improved the discussions on the characterization of sources, chemistry, and concentrations of ambient sulfur oxides.

Section 2.2 – Anthropogenic and Natural Sources of Sulfur Dioxide

This section does a good job of describing the main anthropogenic and natural sources of SO_2 emissions. Figure 2-5 should be updated to include 2012-2015 or 2012-2016 emissions by sector.

Section 2.5 – Environmental Concentrations

Table 2-6 includes "5-min hourly max" and "1-h avg" for 2013, 2014, and 2015. In the second to last column, the 2013 "1-h avg" Max (2,071.0 ppb) is greater than the 2013 "5-min hourly max" Max (1,441.4 ppb) and the 2015 "1-h avg" Max (1,779.0 ppb) is greater than the 2015 "5-min hourly max" Max (1,678 ppb). Please explain how the "1-h avg" Max can be greater than the "5-min hourly max" Max.

Figures 2-13 to 2-18. The SO_2 concentration scale (ppb) in the legend should make the first break point at 75 ppb (3 to 75 ppb) rather than 100 ppb since the current level of the SO_2 NAAQS is 75 ppb. The second break point should be 150 ppb (double the level of the NAAQS), the third break point should be 225 ppb (triple the level of the NAAQS), and the fourth break point should be 300 ppb (quadruple the level of the NAAQS). Also, these break point will allow the figures to show more variability across the maps.

Page 2-54. CAIR has been vacated by the courts and replaced with EPA's Cross-State Air Pollution Rule (CSAPR).

Section 2.5.4 discusses the relationship between hourly mean and peak concentrations. This section states "PMRs were used extensively in the previous SO_2 NAAQS review to evaluate the distribution of 5-minute hourly max concentrations corresponding to a given 1-h avg SO_2 concentration." However,

this section only includes 1.5 pages of text and 2 figures on this topic. Although it not clearly stated, I assume the same number of "5-min hourly max" and "1-h avg" data points used in Table 2-6 (9,149,724 data points from 2013-2015) are plotted in Figure 2-26. However, very little statistical analysis was presented in this section: "Median PMRs obtained from comparing the 5-minute hourly max with the 1-h avg AQS data at sites where both measures were available simultaneously, and neglecting concentrations below 0 ppb, had a range of 1 to 5.5 with a median of 1.3, in reasonable agreement with the predicted range of 1 to 5.4 for the PMR." In addition, this section should look PMR values associated with various percentiles (e.g., 50%, 75%, 90%, 95%, 98%, 99%). Also, the percentage of data points above/below a PMR = 2.67 should be presented since this is the value associated with converting 200 ppb (5-min average) to 75 ppb (1-hour average).

Section 2.6 – Atmospheric Modeling

This section does a much better job of describing the available models and their strengths, weaknesses, and latest updates (especially AERMOD) compared to the previous version of the document. I was pleased to see that the document discussed the differences between models used for regulatory compliance assessments (e.g., related to the 1-h daily max SO_2 standard) and dispersion modeling used in support of health studies where the model must capture concentrations at specified locations and time periods.

Model performance was discussed for a number of historical modeling exercises. It seems that modeling results within a factor of 2 is considered "good". However, biases in the models can have significant impacts on health studies. Whenever possible, model results should be compared to observations and the biases documents. In addition, the modeling results can be adjusted up/down based on model biases identified by comparison to observations to give more realistic spatial and temporal estimations of ambient concentrations.

Dr. Judith Chow

The revised Chapter 2 is well written; it resolves many previous inconsistencies and incorporates most of the panel members' comments on the first draft ISA for SO_x (U.S. EPA, 2015). Following are some discrepancies that need to be addressed:

Section 2.2 Anthropogenic and Natural Sources of Sulfur Dioxide

U.S. Anthropogenic Sources: SO_2 emissions are a rapidly moving target, especially for electric generating units (EGUs), as illustrated in the attached Figure 1 (a version of which should replace Figure 2-5 [Page 2-10] as more relevant to the current situation). The attached Table 1 shows substantial differences relative to the ISA's Table 2-1 (Page 2-9). A modified Figure 2-5 might also include the 2006 on-road and 2007 off-road diesel sulfur fuel standards.

Section 2.5 Environmental Concentrations

Median (50th percentile) concentrations in Figure 2-23 for Pittsburgh (Page 2-58) is a straight line and nearly straight line (except for hour 3:00) for NY in Figure 2-24 (Page 2-59). Section 2.5.1 noted that negative values in the AQS database were excluded in the analysis but concentrations below the lower detection limits (LDLs) were included (Page 2-32, Lines 6-8). As SO_2 concentrations varied diurnally, why are these straight lines at the concentration level around 1 ppb, about 50% of the LDLs of 2 ppb? For Figures 2-23 to 2-25 (Pages 2-58 to 2-61), a footnote should be included that specify how the mean (or arithmetic average) was calculated when many measurements were below the LDLs.

Section 2.6 Dispersion Modeling

A major limitation of AERMOD modeling is the prediction of hourly SO_2 concentrations that don't capture the nature of plume looping and short-duration touchdown. The EPA should require that local and state agencies report all 5-minute measurements, thereby allowing comparison between atmospheric observations and model performance in order to reduce model uncertainties.

Minor Editing:

Page 2-3, Line 9: Figure 2-11 is noted before Figure 2-3, it may be better to refer to Section 2.5.3 instead of Figure 2-11.

Page 2-8 for Figure 2-4: Is this graph representing 377 emission sources that emit more than 2,000 tons per year of SO_2? If so, it should be noted.

Pages 2-56 to 2-61 for Figures 2-22 to 2-25: Label the x-axis and add an explanation for the box plot and legend to Figures 2-22 to 2-24. Note that the legend was included in Figure 2-25 (Page 2-61). For Figure 2-23 through 2-25, are these local standard times as noted on Page 2-57 (Lines 12-14)? Please specify. Also, note that the labels for each hour are not legible.

Pages 2-23 and 2-24: LDL should be consistently defined. On Page 2-23, Line 32, it is defined as "lower detection limit" while on Page 2-24, Table 2-3, it is defined as "lower detectable limit."

References

U.S. EPA, (2015). Integrated Science Assessment for Sulfur Oxides–Health Criteria (External Review Draft). EPA/600/R-15/066, National Center for Environmental Assessment, Office of Research and Development. Research Triangle Park, NC, November 2015.

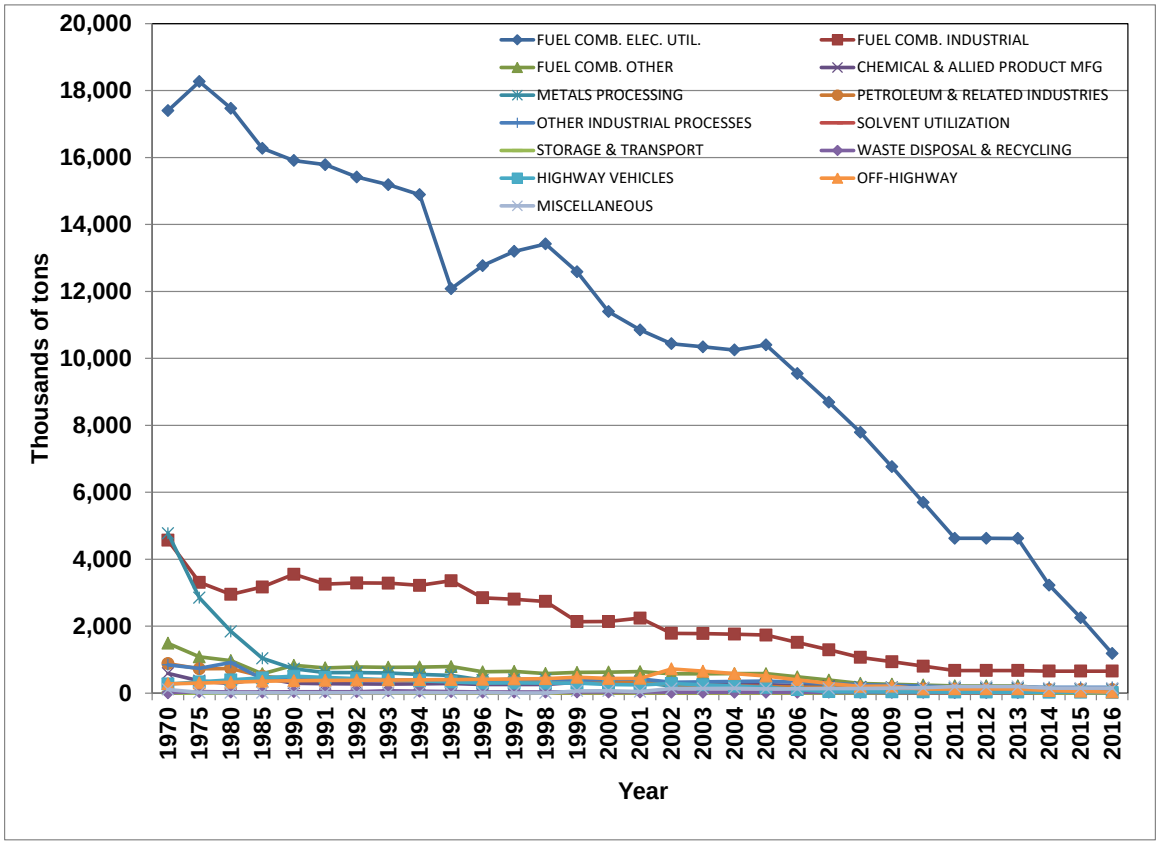

Figure 1. SO_2 emission trends through 2016 from the spreadsheet available at https://www.epa.gov/air-emissions-inventories/air-pollutant-emissions-trends-data. The difference between ISA Figure 2-5 that ends in 2010 and the 2016 emission estimates for EGUs is substantial.

Table 1. Changes in U.S. SO_2 emissions between 2001 and 2016.

Source Category	2016 SO_2 Emissions (x1000 tons)	% of Total Emissions (2016)	2001 SO_2 Emissions (x1000 tons)	Percentage of 2016 to 2001 emissions
FUEL COMB. ELEC. UTIL.	1,185	43.76%	10,850	10.92%
FUEL COMB. INDUSTRIAL	657	24.25%	2,243	29.29%
FUEL COMB. OTHER	131	4.83%	642	20.39%
CHEMICAL & ALLIED PRODUCT MFG	115	4.25%	342	33.60%
METALS PROCESSING	107	3.97%	332	32.40%
PETROLEUM & RELATED INDUSTRIES	103	3.79%	319	32.18%
OTHER INDUSTRIAL PROCESSES	172	6.37%	429	40.19%
SOLVENT UTILIZATION	0	0.00%	1	5.84%
STORAGE & TRANSPORT	3	0.13%	7	51.86%
WASTE DISPOSAL & RECYCLING	37	1.35%	35	105.90%
HIGHWAY VEHICLES	18	0.68%	248	7.43%
OFF-HIGHWAY	31	1.13%	440	6.96%
MISCELLANEOUS	149	5.50%	44	336.47%
Total	2,709		15,932	

Dr. Aaron Cohen

Comments on Chapter 5: Integrated Health Effects of Exposure to Sulfur Oxides

General Comments

The authors have done an impressive job, summarizing in a generally clear and efficient manner a large and complex literature and its evolution since the previous ISA. The links to Chapter 3 (exposure) and Chapter 4 (dosimetry) are very helpful and serve well to buttress arguments made concerning the health evidence.

The authors have largely addressed concerns that I expressed in my comments on the previous draft and I now concur with the causal determinations they have made with regard to the eight classes of health outcomes.

Section 5.2.1 - Respiratory effects - Short-term Exposure

This section effectively presents the evidence for a causal effect based on the evidence for the exacerbation of asthma from both observational and experimental studies. The key role of experimental studies, which provides clear evidence of an effect of SO_2 per se, is well described, and the crosswalk between the animal and human evidence with regard to lags and levels of exposure, the dosimetry and mode of action (Chapter 4) is compelling.

Page 5-6, line 29: Why “In contrast…”?

Page 5-35, line 4: Why are the controlled human exposure studies not interpretable as effects of SO_2 per se?

Page 5-47, line 9-12: This seems a bit of a reach.

Page 5-35, line 16: I could find no previous discussion of “potential differential exposure error.”

Page 5-39, line 1-5: Significance of stratum-specific estimates is not the issue: was there evidence of a trend in effects?

Page 5-39, line 23: Why “In contrast?” Tunnicliffe et al. (2003) seems to corroborate Linn et al. (1983b).

Pages 5-65 to 5-66 and Page 5-71, lines 33-36: **Concentration-response relationship** Although I think the authors have presented the available studies well, I still find the treatment of this issue is overly confident with regard to linearity given the very limited empirical exploration of this issue. More definitive results will likely require much larger studies and more comprehensive and flexible exploration alternatives to linearity

Section 5.2.2 - Respiratory effects - Long-term Exposure

The determination that the evidence is now "suggestive" of a causal relationship rather than "inadequate," rests on two new studies of asthma incidence in children and several experiments in rodents. The experimental studies are key because they serve to mitigate to some extent considerable concerns with attributing the observed effects to SO_2 per se, as opposed, for example, to sulfate particles, given that neither epidemiologic study provided information on co-exposures. As the authors note, Chapter 4 also provides some mechanistic support for this determination.

The authors have, to some extent, addressed issues raised in the review of the previous draft regarding confounder control in the two longitudinal studies of asthma incidence, but an explicit discussion of which risk factors in addition to air pollution are most important to assess regarding confounding and effect measure modification, and what surrogates for them are effective, would still be helpful.

Section 5.3.1 - Cardiovascular effects - Short-Term Exposure

The revised section provides a more critical and balanced assessment of the available evidence, including more critical assessment of the potential for confounding by co-pollutants, and I concur with its determination that despite the accrual of many new studies since the last ISA, the strength of the evidence remains "inadequate." The contrast with the evidence re. short-term exposure on respiratory effects, especially the relative lack of support from experimental studies is telling.

Chinese studies have tended to observe effects on cardiovascular outcomes of short-term exposure to SO_2, e.g., on hospital admissions and emergency department (ED) visits (Page 5-194, lines 7-13) and mortality (Page 5-198, lines 4-11). Given the dominant role of coal burning in air pollution exposure and disease burden in China, it would be critical to control for the most relevant pollutants, i.e., $PM_{2.5}$ or, better yet, sulfate PM, which has been done in few if any Chinese studies.

Section 5.3.2 - Cardiovascular effects - Long-term Exposure

The determination that the evidence regarding cardiovascular effects of long-term exposure to SO_2 is "inadequate" to make causal determination is well-supported by the evidence reviewed in this section.

Page 5-291, lines 9-12 and Table 5-32: The SO_2 estimate for MI incidence from Lipsett et al., 1.98 (0.07-60), is so imprecise that it contributes virtually no information.

Table 5-33 and Page 5-22, lines 24-31: same comment as above with respect to stroke.

Page 5-221, lines 6-9: Without consideration of $PM_{2.5}$ or sulfate PM these Chinese estimates cannot be interpreted as effects of SO_2.

Section 5.3.2.4: Are these longitudinal or cross-sectional studies? Presumably the latter, but if not, more information is needed on cohort follow-up (numbers of repeat measurements, over how many years, loss to follow-up, etc.)

General comment on the Tables (5-32 to 5-34): Please provide the number of events (e.g., MI, stoke, hypertension, etc.) in addition to cohort size.

<u>Section 5.4 - Reproductive and Developmental Effects</u>

The determination that the evidence effects is "inadequate" to make causal determination is well-supported by the evidence reviewed in this section. There is no reason to attribute the observed effects of exposure to SO_2 per se, given that the role of co-pollutants has generally not been addressed in the epidemiologic studies (which also report effects of PM, CO and NO_2 on adverse reproductive outcomes), and the lack of experimental evidence which would support such an association.

<u>Section 5.5.1 - Total mortality - Short-term exposure</u>

The revised chapter provides stronger support for the determination that the evidence remains "suggestive" of a causal relationship and documents more clearly that despite the accrual of more studies, the shortcomings of the literature identified in the 2008 ISA have largely not been addressed. Specifically, the potential for confounding of observed SO_2 effects by co-pollutants remains largely unresolved and, to the extent that confounding by co-pollutants has been assessed, neither $PM_{2.5}$ nor sulfate PM has been included, a particular issue in Chinese studies where coal burning is A, if not *the*, major source of pollution.

The authors correctly note that the exploration of exposure-response relationships has been "limited," but still argue that a log-linear model best describes the relationship between short-term exposure and mortality (Page 5-274, lines 12-21). However, neither Figure 5-24 nor Figure 5-25 appear to support this view.

Section 5.5 - Total mortality - Long-term exposure

Despite the accrual of new studies that observe effects of long-term exposure on mortality, the determination that the evidence regarding mortality and long-term exposure to SO_2 is "inadequate" to make causal determination is generally well-supported by the evidence reviewed in this section.

Page 5-292, lines 27-30: This seems to be the "bottom line." I would consider leading with it.

Table 5-42: Please give the number of deaths (all-cause and cause-specific) and total size of the cohort at entry.

Page 5-291, lines 24-31: These are issues with regard to generalizability as opposed to internal validity and probably do not belong here.

Section 5.5 - Cancer - Long-term exposure

The determination that the evidence regarding cancer and long-term exposure to SO_2 is "inadequate" to make causal determination is well-supported by the evidence reviewed in this section, including the conclusions of other authoritative sources and expert groups.

The authors should note specifically that two of the three largest studies accrued since the last ISA (Krewski et al. 2009 and Brunekreef et al. 2009) report null results for lung cancer. A summary table and/or forest plot would be helpful in presenting the epidemiologic studies.

Dr. Alison C. Cullen

Comments on Chapter 3 - Exposure to Ambient Sulfur Dioxide

EPA has strengthened this chapter since the last draft.

- The consolidation of the copollutant correlation data and content into this chapter helps the overall presentation and document. Comments and questions below are intended to help further refine these sections.
- Improvements to the language about metrics and terminology are helpful, with some additional suggestions below about remaining issues.

Specific points and comments are presented below.

Section 3.2.1 Exposure Metrics

This is a very important section and makes a strong start at distinguishing and defining terms. Additional clarification of terms such as surrogate is needed to help ensure consistent usage across the whole chapter. It is not clear in places whether the goal is to delineate a difference between estimating exposure (such as with models) and representing exposure (such as with measurements other than personal exposure) or instead to group these under the surrogate label.

A tabular presentation similar to Table 3-1 for terms such as surrogate, exposure metric, etc. and the types of error associated with each would help to clarify.

Section 3.2.2 Conceptual Model of Personal Exposure

In Eq 3-1 there is a mixed indexing for the integral. It is necessary to clarify whether integration is across j (the microenvironments) or to integrate across t, in which case, the equation is estimating E_j (rather than E_T).

Pg 3-4 in lines 20-22, Please clarify when Eq 3-5 is assumed in contrast with Ca alone – as currently written the interpretation could be that both are used simultaneously in some contexts.

Section 3.3 Methodological Considerations for Use of Exposure Data

In the opening paragraph, the measurements of microenvironmental concentrations are clearly designated as surrogates for personal SO2 exposure. And in Section 3.3.2 the modeled concentrations are designated as contributing to surrogate estimates. The sentence that follows in line 33 could use

clarification however as it mentions factors such as time-activity patterns and indoor concentrations of SO2 in microenvironments that are incorporated in a portion of the modeling.

Section 3.4 Exposure Assessment, Error and Epidemiological Inference

Table 3-1 is a very nice format for laying out each of the measurement and model exposure assignment method approaches, along with strengths and limitations. Accuracy and precision are conflated in places though, in the Errors and Uncertainties column. Bias (or systematic error) reduces accuracy. Failing to measure where the population exposures occur would lead to bias in the measurements, as would generalizing from one location to another. Precision on the other hand is related to low levels of random error and scatter. These need separate treatment.

The further development of α (first introduced in Eq 3-6), at the top of pg 3-24 (i.e., the slope of the regression between 24 hour personal SO2 concentration and ambient concentration measurements) states the assumption of no nonambient sources of SO2. This raises an overall point that a clarifying discussion of the sources of SO2 would be a useful addition.

Section 3.4.2 Factors Contributing to Error in Estimating Exposure to Ambient SO2

Please clarify whether the phrase "uncertainty in the metric used to represent exposure is a source of exposure error" refers to uncertainty *about which metric is selected* or uncertainty in the chosen metric itself.

Section 3.4.2.1 Activity Patterns

This section has very nice coverage of the relationship between age and activity pattern and also EPA's database (CHAD). Given that disease status is one characteristic of the population of concern for SO2 exposure, another important addition would be a discussion of the influence of disease status on activity patterns or at a minimum an acknowledgement of disease status as a factor, a statement about what data could inform a further interpretation.

Section 3.4.2.2 Spatial Variability

The discussion of Strickland et al (2011) and the importance of using monitors from across the city, rather than a central site monitor, for improving capture of spatial heterogeneity, should be strengthened by the inclusion of a quantitative statement about the size of change in magnitude of observed associations.

Section 3.4.2.3 Temporal Variability

In the Guay et al. (2011) study the population did not move often, an important factor in a long term analysis. It would be helpful to comment in the text on the applicability of this characteristic of the Canadian population, as reflecting the distribution of the population between urban and rural settings, the difference between the US and Canada populations, etc.

Section 3.4.3 Copollutant Relationships

The consolidation of this material into Chapter 3 strengthens the overall document and improves the flow of content. Correlation among copollutants is an important consideration and the detail of these sections is very helpful.

Figures 3-4, 3-5, 3-6, 3-7 figure captions should include a statement of n, the number of monitors represented in the box plot. Also, the inter-quartile range contains a typo in each caption ("inner-quartile).

Pg 3-43 lines 9-11 Clarify the language about higher correlations associated with some sites, this is very general at present. It is true that these sites may introduce greater confounding.

Pg 3-43 Lines 13-16 A statement of the relative magnitude of contribution of various SO2 sources is needed in order to interpret the language about the importance of a change in sulfur standards in 2006.

Pg 3-43 lines 32-34 Is the intent to refer to confounding problems that specifically plague "cleaner sites", clarify either way.

Pg 3-45 line 13 With such high levels of below MDL findings one can neither rule correlations between personal and ambient SO2 in or out (i.e., not just out).

Figure 3-8 is very difficult to read, very small fonts and axis labels. And as with previous figures a statement of sample size would be helpful.

Pg 3-47 top The text reports correlations declining over the last two decades for several copollutants however correlation also rose with respect to O3, not sure why this is not mentioned.

Figure 3-9 regression equations are small and difficult to read, complicating reading which line goes with which pollutant.

Pg 3-48 bottom line 5 the word "reported" should perhaps be "measured" or "estimated" to clarify meaning.

Pg 3-49 line 5-6 Please say something further about how confounding should be appropriately "considered".

Pg 3-49 line 11-13 Rewrite this sentence to clarify, it is confusing at present regarding bias and exposure metric with spatial scale.

Pg 3-49 line 15-31 This sections needs editing for clarity. The phrase "exposure metric" seems to refer to concentration, please clarify. In point (1) is uncertainty referring to uncertainty in the metric used to represent exposure or exposure concentration? Clarify phrase "surrogate parameter of interest." Is it the same as the "surrogate target parameter of interest?" Rewrite sentence about classical error. Clarify discussion of error to distinguish error independent of the measurement or of the measured exposure concentration.

Pg 3-50 line 11-12 Is it always the case that standard error will be underestimated in these conditions?

Pg 3-50 line 31 Is this line referring to people's exposure (as population) or individual?

P 3-51 top Need to include some mention of other factors that could influence exposure estimates such as time activity patterns or other human factors.

Pg 3-51 line 23 The reference to "true air pollutant exposure" should be "pollutant exposure concentration" according to definitions earlier.

Pg 3-52 line 12-13 Rewrite for clarity

Pg 3-52 lines 17-23 Are these slight drops statistically significant? Should say something quantitative to interpret.

Pg 3-53 line 31-33 Be careful about conflating CIs with truth in the language.

Pg 3-54 line 20-21 Rewrite the sentence which begins "In this study, a random error term…" it is difficult to interpret as stated.

Pg 3-55 lines 29-31 It seems circular to say "…long-term concentrations may be well correlated with long-term …"

Pg 3-56 line 28 Not what "exploiting asymptotic properties of variability in exposure concentration" is referring to, please clarify.

Pg 3-61 line 5-7 Is the implication that these urban geographic scales are finer – if yes, state this.

Dr. Delbert Eatough

Comments on Chapter 2

A reasonable portion of my comments on Chapter 2 will be based on how well I think EPA has responded to requests made from the Committee for the First External Review Draft. I will emphasize areas where my research experience played a role in making the requests.

I focus on the following:

In the April 15, 2016, letter to Administrator McCarthy we stated,

> "The CASAC finds that the source categories and definitions of major sources are inconsistent throughout Chapter 2 as well as the entire ISA and recommends that these be consistent. The chapter should include locations and emissions for point sources (energy-generating units, integrated steel and iron mills and smelters) near urban centers."

In the 03/10/16 Draft Report we further stated,

> "The importance of pollution sources and formation of non-sulfate compounds such as inorganic particulate S(IV) species, organic S(IV) species (e.g., bis-hydroxy dimethyl sulfone) and organic S(VI) species (e.g., alkyl sulfates) requires additional discussion. Studies such as Alarie et al. (1973) and Amdur (1971) demonstrated the relationship between exposure to inorganic S(IV) compounds and exacerbation of SO_2 inhalation responses in animals. These compounds are potential confounders or moderators of SO_2 health effects in epidemiological studies where copper smelter or integrated steel mill emissions are abundant and the possible influence of these compounds should be discussed."

And in my Final Comments on Draft Integrated Review Plan (IRP) I outlined in detail what was known about the above outlined chemistry and recommended,

> "Probably a more fruitful set of data to evaluate the relative importance of aerosol S(IV) species associated with smelter emissions would involve past epidemiological studies from about two to three decades ago when smelter emission were much more significant, for example from the TX smelters in El Paseo (ASARCO Cu smelter, closed in 1999), and Corpus Christi (ASARCO Pb smelter, closed in 1985), AZ smelters (ASARCO Cu smelter in Hayden, currently operating and Phelps Dodge Cu smelter in Douglas, closed in 1987), from the Kennecott Cu smelter in Magna, UT prior to construction of the tall stack, from the Tacoma, WA smelter (American Smelting and

> Refining, a Cu smelter specializing in high As ore refining, closed in 1985), or the smelters in Montana (ASARCO Pb smelter in East Helena, closed in 2001, Anaconda Cu smelter in Anaconda, closed in 1981) and Idaho (Bunker Hill Pb smelter in Kellogg, closed in 1982). I know that several epidemiological studies were conducted at these locations, but I am not familiar with the results of these studies with respect to asthma exacerbation. I recommend that EPA look at this older data to see if an estimate of the relative potency of SO2 and smelter associated aerosol S(IV) species can be determined. There will not be data on the concentrations of S(IV) in the aerosols emitted from these sources, so total particulate exposure would need to be used as a surrogate. The importance of elucidating the effect of these exposures is correctly alluded to in the ISA on Page 4-12, Line 11."

It is recognized that the total emission from both smelters and integrated iron and steel mills have decreased significantly over the past 25 to 30 years. In my mind's eye, the critical importance of responding to the above outlined requests made of EPA are twofold:

1. To determine the extent to which the highest concentrations to which a population is currently exposed is due to emissions from sources which are expected to have high concentrations of particulate inorganic S(IV), e.g. emissions from smelters and integrated iron and steel mills.
2. To determine if the older epidemiological data related to asthma conducted where emissions from smelters and integrated iron and steel mills was important indicate a significantly increased exacerbation of respiratory responses in the presence of both SO_2 and particulate inorganic S(IV) compared to that expected for SO_2 alone. This enhanced exacerbation has been seen in animal toxicological studies as summarized by Amdur (1975) which have been summarized by Collucci (1976) (Table 2-5, page 2-61) showing that the effect of transition metal sulfate salts plus SO_2 is significantly greater than the sum of the effects of the two alone. In these studies, the presence of S(IV) in aerosols formed from the combination of ZnO and SO_2 was shown to contain S(IV) species by ESCA analysis (Amdur 1983).

If the answer to 2. is yes, aerosol inorganic S(IV) does appear to be a significant confounder of SO_2 exposure, then the observations seen in 1. become important with respect to health criteria evaluations. I believe the question is significant enough that EPA should address these questions. They have not yet done so in this Second Draft.

For the above outlined reasons, I first address the lack of response to the above outlined requests of CASAC by EPA in Chapter 2.

The question of identifying emissions from smelters and integrated iron and steel mills.

EPA has not done this except for a few oblique comments. Places where I believe this should be done in detail in Chapter 2 are:

Section 2.1, paragraph beginning on line 17. Why are smelter and steel mill emissions not specifically listed? We had asked for them to be highlighted in the chapter.

Figure 2-1. The letter to the administrator for the last review specifically asked to identify and highlight integrate iron and steel and smelter sources. This is not done in Figure 2-1 or throughout Section 2.2.1.

Section 2.2.1, paragraph beginning on line 1, page 2-4. Why are smelter and steel mill emissions not specifically listed? We had asked for them to be highlighted in the chapter.

Table 2-1. Why are smelter and steel mill emissions not specifically listed. We had asked for them to be highlighted in the chapter.

Figure 2-5. If iron and steel and smelters were included as specific sectors in this plot as requested, the reasonableness of looking at older smelter related epidemiological studies to determine if aerosol S(IV) species could be a confounder would be apparent. My suggestion to do so in my final comments given above has not been addressed. Why?

Figure 2-11. This Figure is very germane to the points being discussed in this section. In connection with our request to highlight iron and steel mills and smelters, it would be useful to know the locations (and nearby probable high emitters) for the dark blue (200 to 400ppb) sites in North Dakota (I assume it may be the Bakken oil and gas well field) and the two sites in Arizona (both of which are smelters location in Gila County, see page 2-34) and for the 75 to 200 ppm sites the three sites in Puerto Rico, the site in Alaska and it looks like about 16 sites in the continental US. I have not included the dark blue sites in Hawaii because I assume they are related to volcanic activity. My interest is how many of the various highest sites are EGU related (my suspicion is few or none because of the release of SO_2 from tall stacks), how many are related to other sources and their types and specifically how many include integrated iron and steel mills and smelters.

This probable source information (or at least the location of the samplers) could be included in a Table for the two highest concentration ranges.

Figure 2-12. The difference between Figure 2-11 and Figure 2-12 is the former breaks the 99th percentile of 1-h daily max sulfur dioxide concentrations into ranges 0f < 75, 75 to 200, 200 – 400 and >400 ppb while the later provides the same information for 24-h average sulfur dioxide concentration in ranges of <15, 15 to 50, 50 to 100 and >100 ppb. While not exactly the same, the sites in the two highest concentrations ranges pretty well mirror each other. A Table giving the location of these highest sites and major nearby sources would also be valuable here and would respond to the request of CASAC.

Page 2-37, sentence beginning on line 5. Please add integrated iron and steel mills to the list of sources.

Figures 2-13 through 2-18, the focus areas. General comments. This set of Figures is particularly germane to the question of identifying emissions from smelters and integrated iron and steel mills. I will first make a general suggestion about all the Figures and then address points where high impacts from smelters and integrated iron and steel mills is evident in the data and this should be clearly pointed out. Each Figure has from 2 to 9 triangles denoting sources emitting 2,000 tons/yr or more. I would, in general assume that the very highest emitters (50,000 to 150,000 tons/yr) are EGUs. But most of the lower emitters are probably not, and appear to be influencing nearby sampling sites which give the highest concentrations for that particulate focus area. I can identify some of these sources, but not all. To respond to the request to look at the influence of smelters and integrated iron and steel mills, please add a Table associated with each figure which identifies the indicated sources in each figure.

The following comments on figures in this section highlight how the identification of the sources would be useful with respect to looking at the importance of emission from smelters and integrated iron and steel mills. I will focus on understanding sampling sites with average concentrations above 50 ppb.

Figure 2-13. Cleveland. There are four sources listed in this Figure. I assume the two larger sources are EGUs with tall stacks. I assume the source in Cleveland itself is the ArcelorMittal integrated iron and steel mill. The highest concentrations given in the figure are for E (85.7 ppb), B (61 ppb) and D (52 ppb). It appears sites B and D are high because they are influenced by the integrated iron and steel mill. I have been unable to identify for sure the emissions source near E in Painesville, but it is a smaller source and does appear to contribute to nearby high SO_2 at E. What is it? If it is the Painesville Electric Plant (which does not appear to have a tall stack) it would be a unique opportunity to include the impact of a smaller EGU without a tall stack in the REA study area analysis.

Figure 2-14. Pittsburgh. The only sampling site above 50 ppb is B (56 ppb). All other are well below 50 ppb. Again, please give a table with id of each of the identified 10 sources. This might shed light on why B is a hot spot (with respect to the 50-ppb level) and the other sampling sites are not.

Figure 2-15. New York City. None of the 10 sampling sites in this figure are close to an average of 50 ppb.

Figure 2-16. St. Louis, MO-IL. The only sampling site in this figure with an annual average above 50 ppb is G (80.4 ppb). All the others are well below 50 ppb. The one significant hot spot (G) is located at Heculaneum, the site of a lead smelter which closed at the end of 2013. Is the nearby source the smelter? You would only have had the smelter running for one of the three years. Was this evident in the data?

Figure 2-17. Houston. None of the sampling sites are above (or close to) 50 ppb. Identifying the types of the three sources given would be useful.

Figure 2-18. Gila County, AZ. All the sampling sites are above 100 ppb. Please id the two sources, they should be the copper smelters near Globe and Hayden. The sites near the smelter at Globe are A (116.1 ppb), C (162 ppb) and D (152 ppb). The site near Hayden is B (282 ppb). These are the only sites in the six focus areas above 100 ppb. This highlights the importance of our request about identifying emissions from smelters.

This still leaves open the question of the remaining sites in Figure 2-11 in the two highest concentrations ranges (not including the >400 ppb sites in Hawaii) and what sources influenced them. It is clear from the above discussion of the figures associated with the focus areas that the most significant contributors to the highest concentrations seen at the sampling sites are very heavily influenced by smelters or integrated iron and steel mills. Does this also hold true for the remaining sites in Figure 2-11?

The entire discussion on focus areas following Figure 2-18 ignores our request to focus on source types, esp. integrated iron and steel mills and smelters to help understand the possible role of particulate S(IV) confounders. The above maps of the six focus areas include one monitor where 99th percentiles 5-minute hourly max is above 200 ppb (B in Gila county) and three above 100 (A, C, D in Gila county). These are emissions from copper smelters. There are 5 monitors where the average is above 50 ppb, G in St. Louis (influenced by a lead smelter for one of the three years), B, D and E in Cleveland (B & D influenced by integrated iron and steel mills and E influenced by a nearby small source which may be a small EGU) and one in Pittsburgh (B). Everything else is well below 50.

Any discussion in response to our questions on confounders would have surely caught this.

Page 2-46, paragraph beginning on line 1. This is a particularly good example of where the role of smelters and integrated iron and steel mills is ignored.

The question of identifying non-sulfate S(IV) and S(IV) formation in the atmosphere.

This request has not been addresses in the ISA. In my previous final comments on the first draft ISA I included a detailed section, requested by other members of the committee, on our research in this area. While other aspects of both the formation of SO_2 in the atmosphere and the conversion of SO_2 other species is well covered, the question of identifying non-sulfate S(IV) and S(VI) gas and particulate species formed from the chemistry of SO_2 is not touched on at all. The HERO summary of references both considered and used and considered and not used contains 16 of the references in my final comments on the last draft ISA. The only one used in the currently ISA was a paper on the rapid conversion of SO_2 in the plume from an oil-fired power plant in a fog bank, compared to when the plume was not in a fog bank. That is referenced in Chapter 2. That study also reported on the formation of gas phase dimethyl sulfate when the plume was not in the fog bank. That chemistry is not mentioned anywhere in Chapter 2.

Specific places where identifying non-sulfate S(IV) and S(IV) formation in the atmosphere should have been mentioned follow.

Section 2.1, line 10. Ignores the presence of dimethylsulfate which has been clearly shown to be present in the gas phase in significant amounts compared to SO_2 and is certainly a sulfur oxide.

Section 2.1, line 12. Why is the formation of particulate phase S(IV) species given no attention here or in Section 2.3?

Section 2.2.4.1, first paragraph. Whether we routinely analyze for inorganic S(IV), demethylsulfate, monomethly sulfuric acid or bis hydroxymethyl sulfone, they are there at concentrations that are a very measurable fraction of SO_2 and sulfate. I tried to give you enough data in my comments that could at least mention them. Are they ignored so you can also ignore the question on confounders?

Section 2.3. There is still no discussion of S(IV) species other that SO_2 and H_2SO_3. In addition, there is no attempt to estimate the possible confounding role of aerosol S(IV) species from historical data where exposure to aerosol S(IV) species would be expected to be high. Hence you can ignore the possibility that exposure to emission from a smelter or steel mill may have a different (higher) response from that predicted by the concentrations of SO_2.

Page 2-21, line 2. A fair amount is known about the absorption of SO_2 and formation of stable inorganic S(IV) species. This was summarized in my final comments on the last draft. This is ignored in this draft.

In Summary with Respect to Emissions from Integrated Iron and Steel Mills and the Potential Importance of Particulate Transition Metal S(IV) Species.

I have outlined above the limited toxicological research (Amdur 1975, Colluci 1976) which indicates the combination of transition metal salts and SO_2 leads to a significant enhancement in the exacerbation of respiratory symptoms in experimental animals. I have also alluded to the research by our group on the development of a titration calorimetry oxidation methods (Hansen 1976) and the use of ESCA analysis (Eatough 1978) to study formation of transition metal S(IV) species in emissions, particularly from smelters and integrated steel mills. Based on early results from the combination of these two observations it has been previously postulated (Colluci and Eatough, 1976)**:**

> "Recent studies have demonstrated that stable S(IV), sulfite, species exists in ambient particulates collected along the Wasatch Front in the Salt Lake Basin. It is postulated that these species may be one of the agents critically responsible for the reported adverse health effects attributed to particulate borne sulfur oxides in the Salt Lake City area for the following reasons:

1. Preliminary results for samples collected from study communities in Utah indicate that these is a good correlation between measured ambient sulfite levels and ambient concentrations found for suspended sulfate.
2. Plausible biological mechanisms based on current scientific knowledge exist to explain the apparent etiology of adverse health effects related to sulfite exposure."

This postulated role for the involvement of particulate inorganic S(IV) on the exacerbation of respiratory health by sulfur oxides has still not been tested. It should be.

Other Comments.

Section 2.3.1. The conventional method for indicating radicals should be used, e.g. for HSO3 and HO2, etc. For example, convention would dictate OH· and not OH. This was pointed out in the last review, has still been ignored and is sloppy.

Section 2.5.1, page 2-32 paragraph beginning on line 1. Excluding negative data but including positive data below the detection limit in analysis will tend to give high results. Were the negative data below the detection limit? If so, why not treat them the same as positive data below the detection limit? If they exceed the detection limit, how often and what was the distribution of their occurrence. Any tags in the data set that might explain why if the negative occurrence is regular and above the detection limit? This could certainly bring the entire data set into question.

How common really was this occurrence in the data set?

Pairwise comparison of monitoring sites at the various focus areas, beginning on page 2-49. While interesting, it is not clear to me if there was an objective for this analysis.

Comments on Chapter 1, Integrative Synthesis of the ISA.

In general, this chapter is well written and informative. A few comments where it might be improved follow.

Page 1-2, bullet beginning line 25. I have outlined in my comments on Chapter 2 one area where the objective of including the role of SO_2 within the broader ambient mixture of pollutants has not been meet. The expected result if the hypothesis outlined there, if correct, would be an underestimation of the exacerbation of asthma.

Page 1-2, Sentence beginning on line 36. I assume the "not" should be stricken.

Page 1-2, Sentence beginning on line 38. As noted in my comments on Chapter 2, the potential importance of particulate S(IV) has not been considered.

Page 1-8, Sentence beginning on line 2 the material in ( ). I would question whether 1-h daily max SO2 concentrations greater than 75 ppb have been seen in emissions from power plants because of the impact for tall stack releases. Such cases are certainly not discernable in the focus areas given in Chapter 2, except for Painesville, OH where the EGU stack is very short. Smelters, however, should be added to the list.

Page 1-13, line 9. Add “to” after “tend”

Page 1-31, line 14. EPA has been asked to look at the potential co-pollutant confounding due to particulate S(IV) species, but has elected not to do so. See comments on Chapter 2. If the hypothesis outlined there is correct, this will result in an underestimation of the exacerbation of asthma when an individual is exposed to emissions from smelters or integrated iron or steel mills.

References

Amdur, M.O., Bayles, J., Ugro, V., Dubriel, M., Underhill, D.W., "Respiratory Response of Guinea Pigs to Sulfuric Acid and Sulfate Salts," Presented at the Symposium on Sulfur Pollution and Research Approaches, sponsored by EPA and Duke University (Duke University Medical Center), May 27-28, 1975.

Amdur, MO; McCarthy, JF; Gill, MW (1983) “Effect of Mixing Conditions on Irritant Potency of Zinc Oxide and Sulfur Dioxide,” American Industrial Hygiene Association Journal 44:7-13.

Colucci A.V. (1976) “Sulfur oxides: Current status of knowledge.” Electric Power Research Institute Report EA-316, Palo Alto, California. Access through: http://views.cira.colostate.edu/improve/Publications/EPRI_Reports/EPRI_SOx_Current_Status_Knowledge.pdf

Colucci, A.V. and Eatough D.J. “Determination and Possible Public Health Impact of Transition Metal Sulfite Aerosol Species,” prepared for the Electric Power Research Institute, EC-184, 1976

Eatough D.J., Major T., Ryder J., Hill M., Mangelson N.F., Eatough N.L., Hansen L.D., Meisenheimer R.G. and Fischer J.W. (1978) "The Formation and Stability of Sulfite Species in Aerosols," Atmos. Environ., 12, 263-271.

Hansen, L.D., Whiting, L, Eatough D.J., Jensen T.E., Izatt R.M. (1976) “Determination of Sulfur(IV) and Sulfate in Aerosols by Thermometric Methods,” Anal Chem. 48:634-638.

Dr. H. Christopher Frey

Comments on Chapter 3

The main comments on this chapter are:

- More context regarding the key SO_2 emission sources would be helpful.
- "Microenvironmental modeling" should be treated more consistently throughout the chapter.
- Measurement approaches should also include microenvironmental measurements.
- The summary of the various modeling methods would be improved with a table that compares the modeling approaches.
- The discussion of "errors and uncertainties" in Table 3-1 does not sufficiently distinguish between sources of bias and sources of imprecision, and typically conflates these. In several places, strengths and limitations could be more accurately stated.
- Discussions of air exchange rate, penetration factor, and deposition rate can be better organized, mainly using more accurate subsection headers with some adjustments to the text.
- Klepeis et al. (1996, 2001) are cited. These are good papers, but the reader wonders if there are more recent data. For example, couldn't these points be supported based on the latest version of CHAD?
- The discussion of factors affecting vehicle in-cabin exposure concentrations can be put into more accurate perspective. Several of the factors mentioned are conditional on other factors in terms of importance to in-cabin exposure concentration.
- EPA did a nice job of updating and summarizing CHAD in the 2014 final Health Risk and Exposure Assessment for the Ozone review (EPA-452/R-14-004a). Relevant findings from that work should be mentioned.
- The use of GPS for activity tracking is an ongoing activity at U.S. EPA. Although there are some limitations of using GPS (such as from cell phones), some of the limitations mentioned or alluded to have been or are being addressed in research at U.S. EPA by Dr. Michael Breen. Should update this section to take into account the latest findings from that work.

What follows are detailed point-by-point comments.

Page 3-2, line 36 and after, "Exposure Error" is defined as error with using "concentration metrics" to represent actual exposure. The definition here is unclear. What is an "exposure metric?" How is an exposure metric different from a "surrogate?" A conceptual diagram may help to clarify these. Is "metric" meant to encompass either an unbiased estimate (e.g., personal exposure) or a biased estimate

(e.g., surrogate such as ambient exposure at a central site monitor). As a matter of terminology, not all monitors are centrally located. The term "fixed site monitor" may be more accurately and descriptive of a regulator ambient monitor at a fixed location.

Page 3-3, Equation 3-1: to be more accurate, this would have to be summed over all microenvironments. Alternatively, E_T should be replaced with E_j.

Page 3-4, line 20… this text is confusing. If an epidemiologic study is based on C_a, then there is no assumed model of individual exposure concentration. Thus, Equation 3-5 would seem to be irrelevant. Perhaps the point that is trying to be made is that individual exposure is linearly related to ambient concentration C_a only if C_a accurately represents ambient concentration in the immediate vicinity of each microenvironment. However, it seems incorrect to imply that epidemiologic studies use Equation 5. They typically use only C_a, except in the case of a panel study.

Page 3-5, line 20. It is no longer true that the "vast majority" of SO2 is emitted by coal-fired EGUs. According to EPA emission trends data, SO2 emissions from fuel combustion for electric utilities (which is mostly from coal-fired power plants) contributed over 70 percent of national SO2 emissions from 2003 to 2013. However, since 2013, the share has dropped from 71 percent to 44 percent in 2016. The recent trend is mainly because of price competition between domestic natural gas and domestic coal. Thus, the statement as made, which implies that it is referring to the present moment, is not true.

Page 3-7, first paragraph. Please specify the averaging time for the measurement (e.g., line 5).

Page 3-7, line 24 – this statement is not clear. "not very sensitive to ambient concentration level" of what? Presumably, this refers to potential interferrents (i.e. other pollutants), and not to SO_2. Please clarify by being more specific.

Page 3-7, end of section 3.3.1: missing from this section is a discussion of microenvironmental monitoring. Microenvironmental monitoring comes up later in this chapter. The idea of such monitoring is to obtain a representative measurement of the concentration in an individual microenvironment. In the case of an indoor microenvironment, it is also desirable to simultaneously sample the nearby outdoor environment. Comparisons between outdoor locations, such as outside a residence versus a fixed site monitor, can also be useful. Microenvironmental measurements can use larger instruments than would be used in personal measurements, typically with better accuracy and precision.

Page 3-7, line 30. The header for this section is "Modeling," and later it is revealed that the scope of this section includes stochastic population-based exposure modeling (also referred to in the ISA as microenvironmental modeling). However, the first sentence of this section is only applicable to models that predict outdoor concentration. As another example, the sentence starting on line 33 states that models do not estimate exposures directly, but in fact this is the purpose of stochastic population-based

models. Such models account for time-location patterns and indoor concentrations in various microenvironments. Thus, there is a mismatch between this introduction and the content of the section. The introduction paragraph should be rewritten to more accurately introduce the scope of this section.

Page 3-8, line 14 and related parts of this paragraph. This text would be more internally consistent if the basic physical concept were mentioned first, followed by discussions of how the physical reality is represented in the model and the limitations of the representation. For example, it is not true that average SO_2 concentration always decreases with distance from the source. This is only true from the point of maximum ground level concentration and farther. One can be immediately next to a tall stack and have no exposure to the plume emitted hundreds of feet above until one walks away from the stack to the point of plume "touch down", and one can continue to walk away from the stack until reaching the point of maximum concentration (e.g., under the plume centerline). SPMs appear to disregard the relationship between distance from source and ambient SO_2 concentration near the source, and only apply to ambient SO_2 concentrations past the point of plume touchdown or past the point of maximum ground level concentration. The text vaguely refers to "the stack height issue" (line 17) without explaining what this is about.

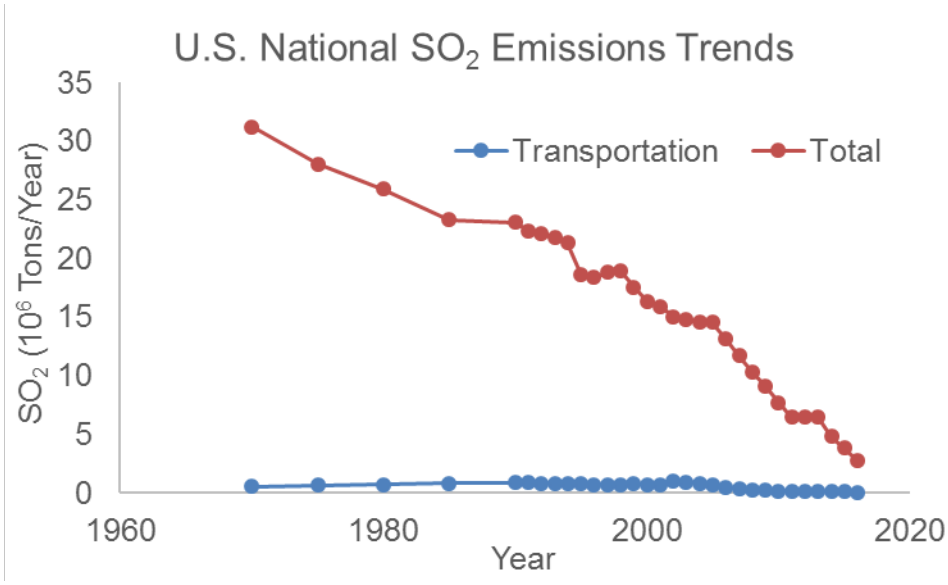

Figure 1. Annual National SO2 Emissions versus Year for Transportation and Total

Page 3-8, line 23. As a matter of context, it should be mentioned that motor vehicles contribute very little to the national SO_2 emission inventory. For example, as shown in Figure 1, emissions from onroad and nonroad transportation combined contribute an average of 3 percent to the total U.S. national emissions since 1970, and in recent years (since 2010) have contributed only 2 percent. The highest annual contribution was 7 percent, in 2002. It can be the case that the relative contribution of transportation sources may be larger in a given region or urban area. The sentence on lines 22-25 is

confusing. Why would a "decrease" in SO_2 be expected "near a highway" – did the authors mean to say that there was not a significant difference in SO_2 in proximity to a highway versus elsewhere in the study region?

Page 3-8, line 27. It is not true that none of these factors are "considered". SPMs account for distance from the source. Distance from the source is a factor taken into account in dispersion models and in CTMs. This should be revised to "none of the factors other than distance from the source…"

Page 3-8, line 29, and other places in the document. Please avoid the word "considered." This is vague and ambiguous. Something can be "considered" by thinking about it. This does not mean it was taken into account quantitatively. To be more specific and clear, here "are considered" would be "are quantified." It may be worth adding, in the context of dispersion modeling, and even to some extent in the context of CTMs, that source characteristics such as release height and other stack parameters are quantifiable.

Page 3-9, line 4. It will be less confusing to break this into two sentences. Delete "and found that" and break here.

Page 3-9, lines 11-12: It is not clear from the description above as to how it is known that EWPM was more accurate. Were both EWPM and SPM validated by comparison to independent ambient monitoring data not used to calibrate the models?

Just before 3.2.2.2 – Inverse distance weighting seems more similar to SPMs than to LURs. Thus, the section on IDW may better fit after SPM and before LUR.

Page 3-9, section 3.2.2.2 – It would help to have a Table that compares the modeling approaches. Table 3-1 compares the approaches from the perspective of epidemiology. However, a table that compares the modeling approaches based on their characteristics would be helpful to the reader. It could be as simple as the suggestion below.

Factors	Type of Model					
	SPM	IDW	LUR	Dispersion	CTM	Microenvironmental
Distance from source	X	X	X	X	X	SO2 Concentration can be estimated based on any of the models to the left or based on monitoring data
Emission Rate		X	x	X	X	
Terrain or Land-Use			X	X	x	
Dispersion				X	x	
Chemistry				x	X	
Human Activity						X
Inhalation						X

The factors above are suggestions – this could be refined.

It may help to add some text to the section on each model that compares and contrasts the modeling approach to those described in previous sections, and to point out similarities. For example, LUR, IDW, and SPM all account for distance from significant point sources.

Page 3-9, line 22-26. Please rewrite this. Very hard to follow. In particular, the idea that a framework "could occur" is very passive and hard to figure out.

Page 3-10, lines 7-10 – hard to parse this. Perhaps break into sentences. What exactly was the improvement? This needs to be clear before stating the change in R2.

Page 3-10, lines 23-25. Not clear as to the point or significance here. Is there more skewness in SO2 concentration than the NO_2 concentration? What is the reason for the difference mentioned here?

Page 3-10, line 29, "predicted" rather than "offered."

Page 3-11, lines 17-24. Hard to parse this. Last sentence is run-on. Just hard to follow this and it should be rewritten for clarity. Similarly, lines 25-26… unclear about background "model" and urban "model." What is meant here by "background?" Seems to be undefined.

Page 3-12, line 16… delete "However," and start sentence.

Page 3-13, lines 4-8. Were the concentrations measured at a fixed site monitor? The correlation with "vehicle sources" is a bit surprising even in 2000 (highway vehicles accounted for only 260,000 tons out of 16,347,000 tons of SO2 emitted nationally).

Page 3-15, lines 22-23 – the way this is written, it implies that wooden structures are problematic when in fact the opposite is true. It is implied here that wooden structures are an example of a "difficulty." In reality, steel frame building pose more of a difficultly because the accuracy of the GPS position is very low for receivers within the building, often to the point of signal loss and inability to report a position.

Page 3-16, lines 3-8. "Exposure Metrics" needed to have been more clearly defined earlier (see comment above). However, another exposure metric is an estimated exposure from a microenvironmental simulation model, or an estimate of outdoor concentration from SPM, IDW, LUR, dispersion, CTM approaches that is used as an exposure surrogate. Unclear as to why the metrics are limited to just measured data. Also, Table 3-1 includes various modeling approaches, but none of them are mentioned in 3.3.3. The text of 3.3.3 should be comparable in scope to the scope of Table 3-1. Also, it is good practice to interpret a table, including its significance, rather than to just mention it.

Table 3-1. For many of the entries on "Errors and Uncertainties" the phrase used is "Potential for bias and reduced precision." The issue here is that bias and reduced precision are not the same thing. Factors that affect bias (systematic error, lack of accuracy) should be treated separately from factors that lead to "reduced" precision (imprecision, random error). For example, it is a source of bias "if the monitor site does not correspond to the location of the exposed population."

Table 3-1 is missing microenvironmental monitors.

Other comments on Table 3-1:

- Central site monitors – this may be better described as fixed site monitors. The use of Central Site Monitors assumes that the relationship between the monitor and outdoor locations at receptors, activity patterns, and infiltration factors are the same when comparing cities. Spatial variation can be estimated, at least partially, if there are multiple monitors in an area.
- Active personal monitors – please indicate the averaging time or the period of time integration of the measurements. Nonambient exposure would lead to bias. Why would it lead to imprecision?
- Passive monitors – high detection limit does not necessarily lead to bias. This depends on how inferences are made based on sample data that include values below the detection limit. There are statistical methods for inferring mean values that are accurate even if data contain non-detects. See, for example:
    - Zhao, Y., and H.C. Frey, "Quantification of Variability and Uncertainty for Censored Data Sets and Application to Air Toxic Emission Factors," Risk Analysis, 24(4):1019-1034 (2004)
    - Frey, H.C., and Y. Zhao, "Quantification of Variability and Uncertainty for Air Toxic Emission Inventories With Censored Emission Factor Data," Environmental Science and Technology, 38(22):6094-6100 (2004).

- Zhao, Y. and H.C. Frey, "Uncertainty for Data with Non-Detects: Air Toxic Emissions from Combustion," Human and Ecological Risk Assessment, 12(6):1171-1191 (Dec 2006)

- Should add a section on microenvironmental concentration monitoring, which can use instruments that are more precise and accurate than those used in personal monitoring.
- Source proximity model: a limitation to mention is that it does not account for emissions, stack parameters, plume dispersion, meteorology, or atmospheric chemistry. In stating limitations, it would help to identify limitations of one method that are addressed by another method in the same table. The discussion of error and uncertainty for SPM is hard to follow – there is too much detail here about "the circle formed….". Given that the SPM does not deal with structural aspects of the physics and chemistry of plume dispersion, the SPM is likely subject to bias moreso than imprecision. There may be imprecision associated with how the model is calibrated – e.g., in terms of a standard error of the estimate, but the bias is much more difficult to quantify.
- If ambient concentration predicted by the model is not the same as at a receptor, please distinguish between bias and imprecision. Bias would imply that the model systematically over- or under-estimates the true concentration. Imprecision would imply that there is random error in the estimate. Given that SPMs do not account for some of the key factors that affect ambient concentration, it seems more likely that bias is the more significant concern here. The model may also be imprecise. The text here is also unclear. Perhaps it is trying to compare the actual and predicted concentrations, but refers to ambient concentration at the receptor (is this the predicted or actual) and the average ambient concentration (where – at the receptor? And is this predicted or actual?).
- For emission weighted proximity models, mention that limitations include 'does not account for stack parameters, plume dispersion, meteorology, or atmospheric chemistry.' The text on errors and uncertainties should be revised – see comments on SPM. Similar to SPM, the text on errors and uncertainties goes into too much detail about the "circle" but not enough distinction of bias versus precision.
- Land use regression: in addition to "land use factors," LUR accounts for distance from a source or a surrogate of source strength. Limitations – to be consistent, the limitations should be stated as "does not account for emission rates, stack parameters, plume dispersion, or atmospheric chemistry, and may account for meteorology only in terms of wind speed or wind direction, depending on model formulation." There should be more discussion either in the text or here as to why spatial resolution affects bias – this is not very clear. If a model is mis-specified, then it is biased. Why is imprecision of concern for a mis-specified model?
- Inverse distance weighting and kriging: see comments above with regard to how to organize some of this information. For example, for limitations, could state, in addition what is mentioned here, that it does not account for emissions, stack parameters, dispersion, and only accounts for

meteorology and chemistry to the extent that it is calibrated to data that have similar meteorology and chemistry. With regard to bias and precision, if sources are not "captured" then it will be biased. Smoothing could also produce bias. While there is potential for negative bias, there is also potential for positive bias – e.g., if the concentration surface does not adequately account for SO_2 deposition or loss processes, or variation in atmospheric/meteorological conditions. Sources of imprecision should be mentioned that are distinct from sources of bias. What factors would lead to imprecision?

- For dispersion modeling, other strengths to mention are that they account for stack parameters, emission rates, mixing height, atmospheric stability, meteorology, terrain complexity.
- For Chemical Transport Models, strengths include accounting for emissions, mixing height, atmospheric stability, meteorology. The limitation is not quite right… local emission sources can be "accounted for" in a large grid cell in terms of being part of the inventory of emissions released into the grid cell, but the model is unable to estimate the ambient concentration associated with the plume from the local source, unless a plume-in-grid model is used.
- For the so-called "microenvironmental model" (which should be referred to really as a stochastic population-based exposure model), another application is to quantify inter-city variability in exposure estimates. The limitation is a bit confusing in that these models simulate synthetic individuals and can estimate inter-individual variability in exposures, but they do not represent actual, specific individuals in a population. The explanation given for errors and uncertainties is about bias, not "reduced precision."

Page 3-16, Section 3.4.1. This section would be better organized if it contained subsections not just on Air Exchange Rate but also on Deposition Rate and Penetration Factor. These three constitute the infiltration factor and should, therefore, receive individual attention.

Page 3-20, Section 3.4.1.1. This section is really about Infiltration Factors and not just Air Exchange Rate. The title should be changed. Klepeis et al. (2001) is a good reference, but it seems dated: isn't it possible to update this estimate based on the current version of CHAD?

Lines 9-10 – list structure here is not parallel.

Page 3-21 – top of page. Should point out that occupant behavior can vary by climate zone and season. Also, building stock, including age, type (multiunit, single unit, etc.) can vary among geographic areas.

Page 3-21, line 9 – does this refer to mean AER? This paragraph could be rewritten for clarity to focus on variability in the mean (or median) AER between seasons and between cities. See also, for example, a similar comparison by Jiao et al. in Table 1 of Environ. Sci. Technol. 2012, 46, 12519−12526.

Page 3-21, line 22 – briefly explain "stack effect."

Page 3-21, line 31: SO2 reaction with indoor sources does not affect AER. This has to do with deposition rate k, which is a separate issue. This is why the header of this section should be retitled. The topic of deposition should start with a new paragraph.

Page 3-22, line 19 – what value(s) of P were used in APEX in the last review?

Page 3-20-31. To put this into perspective, and consider some earlier work:
Ott, W., N. Klepeis, and P. Switzer. Air Change Rates of Motor Vehicles and In-Vehicle Pollutant Concentrations from Secondhand Smoke. Journal of Exposure Analysis and Environmental Epidemiology, Vol. 18, 2008, pp. 312–325.
It is quite clear from Ott et al. (2008) that the most significant factor for vehicle AER is whether a window is open. If a window is open, even partially, the AER is relatively much higher than if all windows are closed, irrespective of HVAC settings, and particularly if the vehicle is moving. The second most important factor is, if the windows are closed, whether the HVAC is recirculating air or taking in fresh air. The fan setting is relatively less influential, as is pointed out. If the vehicle is moving with windows closed and using fresh air intake, then there is sensitivity to vehicle speed. Vehicle age and mileage are correlated, so these are not independent of each other – both are potentially indicators of aging of door seal gaskets that could lead to higher infiltration even if windows are closed. The main point here is that the importance of some of these factors are conditional on other factors… e.g., age of the vehicle doesn't matter much, if at all, if a window is partially open. These factors are not independent of each other.

Page 3-22, line 9 – to be more clear, state that this building was naturally ventilated with open windows, if that was the case, or that there was passive infiltration via the building envelop.

Page 24, line 13: AER, P, and k are also subject to error or uncertainty. Factors such as AER and P depend on building characteristics, climate zone, season.

Section 3.4.2.1: EPA did a nice job of updating and summarizing CHAD in the 2014 final Health Risk and Exposure Assessment for the Ozone review (EPA-452/R-14-004a). One issue that came up in that review was whether activity patterns might be changing over time, especially for children. Since the SO2 review is likely to address similar subpopulations (e.g., school children with asthma), any updates to the 2014 HREA with respect to CHAD would be important to mention in this section. If there are not updates, then it would still be useful to reference the 2014 Ozone HREA and explain (briefly) the key findings that continue to be relevant here.
Are activity patterns different by location/region?

The Klepeis et al. (1996) paper is a good reference but may be out of date. What can be stated based on more recent data (including any relevant more recent estimates based on CHAD from other ISAs or HREAs)?

Table 3-2 – is it possible to also summarize activity level (MET level, ventilation rate) in a similar format? Is Table 3-2 data up-to-date? Confirmed/validated with more recent data?

Page 3-25, lines 11-25: It was a little hard to follow some of this paragraph. It would help if the entire paragraph was focused on high exposures. The statement about "more likely to spend time indoors" seems to hint at low exposure. This could be reworded as white, black, and Hispanic study participants were likely to spend more time outdoors than Chinese participants, to put the focus on factors leading to high exposures (to be consistent with the rest of the paragraph).

Page 3-26, line 24 – it would help to state how many subjects are in CHAD and how many of these have diaries for more than one day.

Page 3-26, line 34: hard to understand this. "frequency of Android-based phones" doesn't make much sense. Perhaps this is trying to refer to the frequency with which GPS coordinate positions changed by more than some distance threshold? Please clarify.

Top of page 3-27. Need more of a story here… GPS signal can be weak or lost when inside a steel frame building. If someone is in a steel frame building, then it is plausible that the positional accuracy could be 342.3 meters. However, in such situations, it is usually known as to what were the last recorded coordinates prior to signal loss or a weak signal and/or the subsequent record coordinates after signal acquisition or signal strengthening. Thus, it may still be possible to infer that the person was inside the building. Dr. Michael Breen at US EPA has been developed MicroTrac for this very purpose. It is surprising that this is not mentioned. GPS signal reception may also depend on what GPS receiver technology is being used.

Page 3-28: lines 10-11: it is more useful that 2/3 of the population lives WITHIN 15 km than that 1/3 live OUTSIDE 15 km? i.e. make a positive rather than negative statement on this point.

Page 3-35, top of page – is it possible that this was also a period of substantial emission reduction?

Page 3-35, line 12 – sentence fragment, delete "although."

Page 3-36, line 5 – is the error related to time of day or just related to the magnitude of ambient concentration?

Page 3-37, line 10 – human clinical data can help with causality inferences.

Page 3-37, lines 11-23 –SO2 concentrations are not well correlated with O3 and PM2.5 concentrations, which can also help in identifying effects specific to SO2.

Page 3-43, line 11- perhaps this linkage may have been significant in some places at some times, but transportation has always contributed far less than 10% of the national SO2 emission inventory, typically 3 percent. It may be possible that the contribution could have been larger in some areas at some times, but some context/justification would help.

Page 3-43, line 30 – some context could be given to this page – there is also a role for clinical controlled human studies to provide evidential support.

Page 3-49: lines 19, 26 – please replace “can be considered as” with “is”

Page 3-51 line 13 – time weighted averaging of what? Wording here is not clear.

Page 3.54, lines 3-4: are the authors trying to say that bias increases with increasing positive correlation? Not sure that this makes sense. Please check.

Page 3-55, lines 21-32: is this about ambient concentration or exposure concentration? Should specify.

Page 3-58, line 5, incomplete comparison “are used much less frequently” than what?

Page 3-60, line 12…. “microenvironmental models” were mentioned earlier in the chapter, but are not mentioned here. Should be consistent.

Page 3-60, line 27 – is this in the context of central site monitors?

Dr. Steven Hanna

Note that my expertise is primarily in atmospheric transport and dispersion modeling and analysis of observed concentrations, and my comments focus on those areas. I was asked to comment on Chapter 2 of the 2nd draft ISA on the areas under "Ambient Air Concentrations."

General comment 1 – I see that two sections of interest to me have been enhanced in response to our comments a year ago. These are section 2.5.4 "Relationships between hourly mean and peak concentration" and section 2.6.1 "Dispersion modeling." These revisions lead to much better justifications for subsequent analysis. However, I have some specific addition suggestions as described below.

General Comment 2 – Although the EPA made the suggested revisions to some sections, they did not put the revised concepts into practice in other major sections describing the general processes of plume dispersion.

General Comment 3 – I realize that this is an internal EPA document, but it is still a scientific study and the concentration analyses and dispersion modeling sections should review the literature outside of the EPA. The document is being made available for "public" review. There is much work that has gone on in other U.S. agencies and across the globe on topics such as statistical analysis of meteorological and air quality data and development and application of dispersion models for all averaging times. I would like to see a comprehensive review done and then pros and cons listed for alternative methods and rationale for choosing a specific method.

As an example of a fundamental reference, see:

EPA, 1974: Proceedings of Symposium on Statistical Aspects of Air Quality data. Chapel Hill NC Nov 1972, EPA-650/4-74-038 Environmental Monitoring Series, 270 pp. https://nepis.epa.gov/Exe/ZyPURL.cgi?Dockey=2000X6RH.txt.

This conference proceedings references has relevant papers by many authors. For example, it contains a paper by Ralph Larsen of the EPA, who suggests the log-normal distribution, and Frank Gifford of the Atmospheric Turbulence and Diffusion Laboratory (ATDL), who proposes a scientific explanation for why the concentration distribution should be log-normal.

Specific Comments

p 2-17 – Section 2.3 on atmospheric chemistry and fate – I see that there are many equations and references in this section, plus detailed discussions and justifications. This format should be used for all sections.
p 2-29, Fig 2-10. It has been stated that there are 438 sampling sites but this figure contains far fewer than 438 dots. This is probably because there are several areas in the US where sampling sites are concentrated, such as around major industrial areas and in metropolitan areas, and so the points fall on top of each other. As suggested earlier, it would be helpful to include zoomed in figures to show some of these concentrated areas.

Section 2.5.2 -Spatial Variability (with many tables and figures) – As commented a year ago, this analysis of spatial variability appears to have been done without reviewing the U.S. research in other agencies and the international literature on this topic. It is known that atmospheric variables are influenced by the full spectrum of motions, which are characterized by random turbulence variations and effective space scales. For example, SCIPUFF parameterizes this space scale of turbulence in its formulations. The mesoscale length scale is a few tens of km. The EPA should also look at its own literature such as the 1970s-1980s RAMS study of SO_2 in the St Louis metropolitan area, where there were many samplers operating and several comprehensive analysis reports published.

Pasquill's book "Atmospheric Diffusion" provides the mathematical basis for the space and time variations of meteorological variables and pollutants.

There have been several attempts to fit basic statistical distribution functions to sets of observations such as these (e.g., the data in Table 2-6 on p 2-35). "Long-tail" distributions are usually found, and can be fit by, for example, a log-normal distribution, or an exponential distribution that accounts for intermittency. See the EPA report cited above. The Robust High Concentration Method used by OAQPS is based on a specific distribution. As stated in later report sections where the text responds to my earlier comments, the variations are also functions of meteorological parameters (e.g., stability, wind speed) plus nearness to large point sources and the stack height. Those concepts should be tested with these data.

p 2-50 (Fig 2-19 for 1 hr avg) – Here is an example of where the "distance scale" could be estimated, although it would be better to have ln C on the vertical axis. My eyeball estimates of the distance scale are about 50 km for all sites except NY, where 150 km might be better. The scientific reason for the larger scale at NY is that NY is the largest metropolitan area. Plus NY is in the middle of the northeast US SO_2 polluted area and has input coming from upwind power plants and other urban area.

p 2-52 top par – These conclusions are truisms that would be expected after a review of the literature on the topic.

Section 2.5.3 – Temporal variability – Again, there are no non-EPA references. The topic of time variability would have a much broader literature that the topic of spatial variability since it is easier for researchers to obtain the data and analyze them from one sampler. Ralph Larsen of the EPA spent much time during his career on this topic and his work should be summarized. His suggested distribution functions should be compared with these new results.
The conclusions about seasonal variability etc, should be compared with those of the many similar analyses.

Section 2.5.3.3 - Define "diel" in the section caption. Figs 2-23 and 2-24 – I would help my comparison of sites if the same vertical scale were used.

Section 2.5.4 – Thank you for adding the expanded references and discussions to this subsection.

p 2-62, line 8 – Clarify what exactly is "longer." I think that you are referring to the overall sampling time for the data. The current report is using the 5 min averaged peak C in a 24 hr period and the 1 hr averaged peak C in a 24 hr period. If you expanded the sampling time to one week, the 5 min peak during a week and the 1 hr peak would both likely be larger (they would never be smaller). Also, I would discard the possibility that the 5 min peak equals the 1 hr peak, since that would never happen with good data (high precision and low threshold).

In the scatterplots, there are so many points that they overlap each other and it is not possible where the "best fit line" would pass. Can a best fit line be shown? Note that the maximum possible ratio of peak 5 min C to one hour C is 12 (i.e., the case where the hour contains all zero 5 min averages C's except one). Please include the well-known Turner 1970 Workbook 0.2 power law prediction, which would be a line of $12^{0.2} = 1.64$. And also his later suggestion that the power is 0.5 (which would be a line of $12^{0.5} = 3.5$).

p 2-65 top and Table 2-9 – What variables are used in calculating the correlation coefficient?

Section 2.5.5 – Background concentrations – Don't we know what SO_2 concentrations are occurring at west coast stations on the coast? Don't SO_2 sources in Canada and Mexico cause transport to the US? How exactly is "background" defined? For example, earlier in this report, you defined six regions in the US with dimensions of about 50 to 100 km. Don't each of these regions have a "background" SO_2 concentration on their upwind boundary?

p 2.6.1 Dispersion modeling – This section is much improved (see general comment 1). However, it would help readers' understanding to point out that the basic dispersion modeling theory is not locked into any averaging time. AERMOD uses 1 hr because that was mandated by the Clean Air Act. The basic Taylor, Batchelor, and Pasquill references include averaging time as a variable and can show how concentrations vary with averaging time. European air quality regulatory models such as ADMS allow any averaging time. SCIPUFF (the most widely used dispersion model in the U.S. outside of the EPA) allows any

averaging time. AERMOD's basic boundary layer turbulence parameterizations would also allow smaller averaging times than 1 hr. All that is needed is some modifications to the program statements (I confirmed this two days ago with Drs. Jeff Weil and Akula Venkatram, the two main developers of AERMOD).

p 2-70, first line of last par – Insert "by EPA" after "model." Use this qualifier elsewhere in the section, too (such as lines 3, 4 14, and 29 of p 2-71).

p 2-72 line 36 – Did Weil call these "errors?" Usually natural variability is a large contributor, too.

p 2-73, lines 13-17 – Mention the other agencies who are now using WRF for inputs to dispersion models (NRC, DOE, DOD). This is a worldwide trend as computer power increases. There are found to be a few problems in applications – such as a tendency of WRF (or any NWP model) to have wind direction errors of 30 degrees or more, and a failure of the WRF grid (4 km or 12 km) to account for subgrid terrain effects.

p 2-73 lines 33-34 – The RHC is based on a distribution assumption. Why can't this approach be used in the current report?

p 2-74 and 2-75 – As stated above, AERMOD could be modified to include any averaging time (as done in most other dispersion models).

Dr. Jack Harkema

EPA Charge Questions for Executive Summary and Chapter 1 of the First Draft ISA

1. *Please comment on the clarity with which the Executive Summary communicates the key information from the SOx ISA. Please provide recommendations on information that should be added or information that should be left for discussion in the subsequent chapters of the SOx ISA.*
2. *Please comment on the usefulness and effectiveness of the summary presentation [in Chapter 1]. Please provide recommendations on approaches that may improve the communication of key findings to varied audiences and the synthesis of available information across subject areas. What information should be added or is more appropriate to leave for discussion in the subsequent detailed chapters?*

Executive Summary (ES) and Chapter 1

General Comments

After reviewing the first draft, the Panel had the following specific comments and suggestions for the second draft: 1) It should be clearly stated in the ES and Chapter 1 that the controlled human exposure studies are the principal rationale behind the 2010 1h SO2 NAAQS and that this standard also provides protection from chronic exposure effects; 2) Correlation of maximum 5-minute SO2 concentrations with corresponding 1-hour concentrations should be summarized; 3) Definitions for short- and long-term exposures need to be clear and consistent throughout the text of the entire ISA, including the Executive Summary; 4) Some of the footnotes used in the first page of the Executive Summary should elevated to the body of the text; 5) It is important that ambient background concentrations of SO2 be mentioned in this summary section of the ISA. Overall the authors have adequately addressed all of these suggestions.

The summarized material (and format) in these introductory sections of the ISA appropriately highlights and summarizes the important information provided in the subsequent chapters. Integration of the chapters is somewhat limited but most effectively captured in the sections on causality determinations. The authors' causality determination for the respiratory health effects related to short-term sulfur dioxide (SO2) exposures has been appropriately highlighted in both the ES and Chapter 1, but the one change in causality determination since the last review (ISA 2008; respiratory effects related to long-term exposure) needs to be more clearly stated in the ES.

Summary tables with chapter references have been appropriately used to synthesize and streamline the study results that were used to determine causality. I have only a few remaining specific suggestions listed below for the final ISA draft.

Specific Comments

Though the ES in this draft contains much less technical jargon and phrasing than the previous draft, the authors are still encouraged to further refine this important section of the ISA so that it is even more understandable (readable) for a wider sector of the public.

As requested by the Panel, the authors have clearly and consistently defined what they mean by short-term (minutes up to a month) and long-term (greater than a month to years) exposures to SO2. However, the subtext of all the tables should also include these definitions that are found throughout the text.

Page l, lines 25-34. It is not clear what questions or uncertainties remain that prevent raising the causality determination for long-term SO2 exposures and respiratory health effects. This should be briefly clarified.

Page lii, lines 29-30. A brief example of the evidence for increased health effect risk for older adults should be provided.

Dr. Farla Kaufman

Comments on Chapter 1

The revisions to the Executive Summary provided much improvement. The Chapter reads quite well with less redundancy and is very accessible for a nontechnical audience.

P 1-20 line 11 exposure error s/b exposure measurement error

P 1-29 line 27-28
More recent NHIS data shows somewhat lower prevalence. For females, those under age 15 years (6.8%). For males, those under age 15 years (9.5%). However, for children under age 15 years, the sex-adjusted prevalence of current asthma was higher among non-Hispanic black children (15.4%) compared with Hispanic children (8.1%) and non-Hispanic white children (6.2%).

Comments on Chapter 4

P 4-8 line 30 patters s/b patterns

P 4-28 line 8-13 No mention of mixtures in summary

Comments on Chapter 5

Revisions have improved Chapter 5 considerably. The section on respiratory effects is very well-presented and clear. Other sections are not as coherent or integrated (see comments below). Overall, the chapter is still very lengthy and could be more focused and concise in sections.

The characterization of the evidence and rationale for causal determinations of effects of SO_2 outside the respiratory system is consistent with the EPA's causal framework.

The characterization of respiratory effects observed in controlled human exposure and epidemiologic studies, particularly in different populations and lifestages is adequate.

P 5-17 line 21 concentration s/b concentrations

5-24 line 6- 8 Since the sentences above discuss obese children, it should be made clear in the comparison is between normal-weight school-aged children and normal-weight adolescents and adults.

Page 5-25 line 1-2 does not seem as if this sentence is actually summarizing the information presented in this section.

5-34 line 4 I am not sure that <u>most</u> of the studies did show correlations above 0.5. Could you provide references in brackets for the ones that fall into this category. It is not clear what is being references by "previous studies."

5-36 Table 5-7 should Magnussen be under the "M's"

5-40 line 11 Table 5-2 s/b 5-8

line 20-21 Sentence implies that results from Anyenda et al. are included in the Figure 5S-1, which they are not.

line 22 Please specify which studies are referred to here. Does not seem to be referring to studies mentioned above (Maestrelli et al., Anyenda et al., or Wiwatanadate and Liwsrisakun) as the correlations in these studies were moderately correlated at most, while many were weakly correlated.

Line 23 Which studies are being referred to here since Anyenda did not report PM metrics.

Page 5-44 line 32 Is it less uncertainty or more uncertainty?

5-45 Figure 5-2 There are two separate lines for both Delfino et al. and Segala et al., but no indication of how the two lines for each study differ from each other. Could use at least footnotes for these.

5-48 line 4 I had trouble accessing this Supplemental Table 5S-5

<u>Section 5.4</u>

This section is improved. However, I find that there still could be more discussion and integration of studies in terms of strengths and limitations for studies of outcomes such as preterm birth and fetal growth. Since inconsistent evidence from epidemiologic studies may be due to differences between studies, adequate consideration and integration of the strengths and weaknesses of the studies is required to weigh the strength of the body of evidence (as mentioned in previous comments).

Tables and text could be reviewed for accuracy in this section regarding which studies reported co-pollutant analyses and the adjusted risk estimates (e.g. Liu et al. 2003).

P 5-237 Table 5-37 is not referenced in the text on page 5-239 line 25

Issues Relevant to Chapter 5 and 6

The issue of responders and non-responders was discussed at the March 20-21st meeting. In relation to this subject, in the 1st draft of the ISA it was noted that Rubinstein et al. (1990) reported NO_2 induced greater airway responsiveness in only study subject. However, as I noted in my comments on that draft, there was no indication of what percentage that one subject represented of the study population. It is important to include that one subject was one of nine, >10% of those tested. Although it is not clear what percentage of the general population these subjects do represent, it is consistent with other reported findings of responders and non-responders to various exposures.

Discussion of increased oral breathing in overweight/obese children, especially boys, appears in Chapter 4 (4-10) and Chapter 5 (5-24, 5-34) with the acknowledgment that "school-aged children, particularly boys and perhaps obese children, should be expected to experience greater responsiveness (i.e., larger decrements in lung function) following exposure to SO_2 than normal-weight adolescents and adults." This information should also be carried forward into Chapter 6, as it should also be considered in the REA.

As I mentioned during the March meeting, the section in this Chapter on obesity, as well as any other mention of the subject within this Chapter, seems to have been removed. To my knowledge, there were no comments from the CASAC with regard to removing discussion of this group of overweight/obese individuals as possibly an at-risk population.

Discussion during the March meeting also included recognition that other socio-demographic variables (such as race/ethnicity, gender, poverty status, housing conditions, overweight/obese [as mentioned above], etc.) are factors to be taken into consideration in the REA in generating simulated populations, as the prevalence of asthma can vary with these factors. As such, Chapter 6 should include discussion of these factors with regard to asthma, along with updated prevalence rates.

Dr. Donna Kenski

The charge was to comment on revisions, but since I wasn't on the panel for the first round of ISA reviews, I can't comment on EPA's responsiveness. My comments are instead about the general adequacy of the Atmospheric Chemistry section and its coverage. Overall I find that the EPA staff has done an excellent job summarizing state of the science, both for the most recent data and material covered in the 2009 ISA and prior reviews. Considering the ISA's many authors and topics, it was remarkably easy to read and sounded mostly as if it had been written by one person. I commend the editor(s) who achieved this. My comments below are thus mostly minor requests for clarification or correction.

Comments on Chapter 2

p. 2-2, lines 23-25: Occasionally there is a jarring qualifier in the stream of mostly straightforward text; for example, these lines 23-25 say, "Anthropogenic emissions of sulfur are…emerging from point sources in quantities that MAY substantially affect local and regional air quality". In a 600+ page document devoted to those effects, isn't the 'may' unnecessary?

p. 2-5, line 22; "four sites ~~in~~ located at the Port…" (delete in)

p. 2-10, Fig. 2-5. The 12 categories of emissions can't all be distinguished on this plot in either the printed document or screen version. Please avoid the temptation to squeeze too much information on one plot. The data are too important to obscure in this way. It could be reworked by combining some of the minor source categories and/or using simple lines instead of the stacked/filled lines, which don't really allow the user to compare among the categories accurately. As others have noted, the data should be updated through 2015.

p. 2-16, Table 2-2: What order are these sources listed in? It's almost, but not quite, total sulfur. A rationale for the existing order, or reordering by total S mass would be helpful for making quicker comparisons. Also, the tire combustion numbers look wrong.

p. 2-18, line 19: rates -> rate

p. 2-21, line 6: glyoxyl -> glyoxal

p. 2-22, line 31: remove the extraneous comma after 'and'

p. 2-24, Table 2-3: The specified lag, rise, and fall times of current FRMs are fine for hourly data but less good for 5-minute average data. Presumably EPA and the health community will continue mining the 5-minute data more extensively in the future and filling in some of the gaps that now exist (e.g., reporting all 12 5-min averages each hour rather than just the max). It's not clear to me what the impact of these response times might be on the future data analyses, if any. Please take a closer look, perhaps by comparing collocated 5-min measurements from FRMs to some of the faster-responding instruments described in Section 2.4.2 to see what, if any, differences arise.

p. 2-32, lines 6-7: Please justify the decision to exclude the negative values. They are included in the database because they represent the monitors' response to real atmospheric variability. Eliminating them thus introduces bias to your analysis. That bias may be small and most apparent on the low end of the distribution, but it's still unnecessary unless you have a good reason for eliminating the negative values.

p. 2-56: Figs. 2-22, 2-23, and 2-24 should have the same explanatory note as Fig. 2-25 (or just put it on Fig 2-22, the first of this series of graphs)

Dr. Richard Schlesinger

p. 4.4, l. 16-18.	Bronchoconstriction can be initiated by irritants in the URT as well as those in the LRT
p. 4.8, l. 27.	Proportion should read proportional
p. 4-10, l. 18-20.	It is not clear what is meant by even during exercise since with exercise, there is a switch to even increased mouth breathing
p. 4-10, l. 26-27	It states that the rate and route of breathing both have great effects on the magnitude of SO2 absorption in the URT and penetration of SO2 to the lower airways. However, earlier in the chapter it is stated that, “during normal breathing....95% or greater SO2 absorption occurs in nasal passages...”. These two comments seem to contradict each other.
p. 4-24, l. 17-18	It is not clear how the cited study supports the statement since it was a combination of both SO2 and NO2
p. 4-27 l. 35 - 4.28, l. 1-7	The last sentence indicating that the study provides consistency with SO2 playing a role in the exacerbation of AHR is an overreach, since as noted, the other components of the atmosphere may have contributed to the response. This study should not be cited here as evidence for exacerbation of AHR.
FIG 4-2	Short term effects are ascribed solely to the formation of sulfite in the ELF and the potential effect of direct irritant effect due to formation of H+ is ignored in this paradigm
FIG 4-3	This suggests that similar outcomes and endpoints occur by different key events in long term vs short term exposures. Is that actually the case? The text notes that the effect of long term exposures is due to the formation of reactive products in ELF, which are the same ones following short term exposure. This is confusing.

Dr. Elizabeth A. (Lianne) Sheppard

Comments on Chapter 3

- Overall the new organization represents good progress. I think the new main sections are good choices. However, I am not convinced that the chapter currently meets its overall objective as outlined in Dr. Vandenberg's January letter. There are details and overall perspectives in the various sections that still need work and many of my comments address specific details. While this chapter is improved from the previous version, it is still a difficult chapter to digest. In particular, I think the chapter is still a work in progress in terms of supporting the evaluation of the strength of inference in epi studies in Chapter 5. I suggest EPA take our suggestions for improvement, integrate them as best they can in the revised ISA, and more importantly, consider how to make further improvements in the ISA for the next criteria air pollutant to be reviewed. Ideally this chapter should become better with each new criteria pollutant ISA since a decent fraction of the material that belongs in this chapter applies to all criteria pollutants.
- Section 3.1: I suggest the introduction more clearly lay out the purpose of this chapter. In my opinion this chapter is essential in order to properly interpret the epi studies discussed in Chapter 5. It also needs to clearly distinguish epidemiologic vs. risk assessment applications of exposure models. I believe there should be some attention to exposure for risk assessment in this chapter because the ISA needs to provide all aspects of the scientific context of exposure, both for later chapters of the ISA, as well as for the REA and the PA.
- P. 3-2 lines 5-6: The statement that ambient concentrations are more relevant to epi studies is unclear and potentially misleading. I think the point is that many epi studies that use fixed site monitors or modeled exposures to derive the exposure metric for the health analysis rely exclusively on ambient concentration instead of exposure concentration. More scientifically relevant is personal (total or ambient source) exposure concentration. However, when this is not measured, epi studies use ambient concentration instead of exposure concentration.
- P. 3-2 lines 20-27: This text needs to be improved. The definitions of differential and non-differential error and their link to misclassification need correcting.
    - Make sure throughout the chapter that discussions of surrogate and exposure error are clearly pointing to either 1) the exposure metric itself or 2) the impact it has on the estimate of the target parameter of interest in the epi study. While the text in this section does clearly refer to the latter, I think in this chapter these ideas are not always clearly separated and thus it can be easy to misunderstand the points being made. For instance, in discussing sources and impacts of bias, it matters whether the focus is on estimating the exposure itself or on estimating the target parameter of interest in the epi study. The same point can be made about precision. (See e.g. the comment about line 29 below)

  - It is too simplistic to think of differential error as systematic error. The technical definition is that the mismeasured covariate has information about the outcome beyond what is contained in the true exposure. Both differential and non-differential error are defined in terms of the information contained in the mismeasured exposure covariate after conditioning the outcome on both the true exposure and other covariates in the model not measured with error. With differential error information about the outcome remains; with non-differential error it does not. See Carroll et al. (2006, 2nd edition) *Measurement error in nonlinear models: a modern perspective*, Section 2.5 and Baker et al. (2008) *Environmental Epidemiology: study methods and application*, the Armstrong chapter on measurement error (chapter 5), section 5.2.
  - It is correct that exposure misclassification refers to categorical covariates while exposure measurement error refers to continuous covariates. Often, when speaking generally about the topic, the term exposure measurement error will be used for both. Differential and non-differential refer to both measurement error and misclassification.
- P 3-2 line 29: Some correction is needed. The previous sentence referred to the effect estimate (i.e. the estimate of the target parameter of interest in a disease model for an epi study), but this sentence is defining bias (and precision??) in terms of the exposure data.
- P. 3-2 line 33: The attenuation of effect estimates is characteristic of (pure) classical measurement error, which is a type of nondifferential error. Pure Berkson error doesn't tend to attenuate effect estimates (and it doesn't bias the effect estimates at all when the disease model is linear).
- Pp 3-2 – 3-3, paragraph starting line 36: I appreciate the distinction being attempted regarding measurement error and a different target parameter of interest in the disease model. I'm not sure that I would call this "exposure error" and then go on to define classical and Berkson error. Also, I would drop the reference to the central site monitor measurement (p 3-3 line 7) as it depends on the context as to whether or not I would agree with this being an example of Berkson error.
- Section 3.2.3 (exposure considerations specific to SO2) gives useful perspective and is concise.
- The introduction to Section 3.3 should clarify what is meant by "use" in this section heading. There are three main uses of exposure assessments: to quantify exposure for environmental surveillance and compliance; to quantify exposure for epidemiologic inference about health effects; and to quantify exposure for risk and exposure assessment. Which of these (or something else) is the context for the term "use" in this section?
- Section 3.3.2 (modeling) is still difficult to digest in terms of the goals of this document. I think we want to understand 1) what are different approaches to exposure modeling and 2) to appreciate how these various approaches affect the conclusions to be drawn from the epidemiological studies. The document has made some progress on 1) and could be better developed with respect to 2). (See also my comments on section 3.4.) However, regarding 1), improvements would still promote better understanding. I suggest dividing the discussion into approaches that focus on modeling measurements of SO2 vs those that focus on other ways of modeling ambient SO2 that may not use any measurement data (except perhaps to validate the modeling) but instead rely on physics, chemistry, emissions, and other information to predict concentration. Right now these are intermixed

with the LUR and IDW approaches to modeling measurements appearing in the middle of discussions of SPM, dispersion models and CTMs. Regarding 2), I believe that the important idea is that the predictions from these models are used in epidemiological studies, predominantly cohort studies, as an exposure covariate. So for the latter point it matters whether the model outputs are reasonable for the application. This could be affected by many factors such as data time period and representativeness (e.g. is a LUR based on data collected during one 2-week period useful at all for an epi cohort study focusing on long-term average exposure?) as well as the quality of the model results. There are important distinctions between models that use measurement data directly (i.e. statistical models such as LUR, IDW) and those that rely on other characteristics rather than measurements (i.e. SPM, CTM, dispersion models). Recently developed measurement error correction methods address statistical models only.

- Section 3.3.2.2: The discussion on LUR should say early on that the focus is on a long-term average pollutant measure. LUR models capture spatial variability (as is said), so they are most useful in cohort studies where the focus is on long-term average pollutant exposure. Overall the discussion of the examples does not provide enough context to help readers appreciate how LUR model estimates affect inference in epidemiology studies, or really even to understand the examples themselves at face value. For instance, the averaging time of the measurements is not mentioned in the discussion of any of the examples.
- P 3-9 lines 19-21: LUR models are used to **predict** exposure (concentration) at unmeasured locations. I wouldn't characterize this as "increasing heterogeneity in the spatial resolution". Rather LUR models allow the ability to characterize more completely the spatial variation, by predicting at arbitrary locations. In reality many regression predictions (e.g. from a LUR) will result in smoother surfaces than measurements. The wording suggests the opposite.
- P 3-9 sentence starting line 22: the apparently agnostic list of metrics for characterizing the results from LURs misses some important statistical ideas. The in-sample estimates of R^2 or RMSE are estimates of training error, not generalization error and thus don't tell us what we want to know about how a predictive model performs. Cross-validation is one tool for estimating out of sample performance of a model. Estimates of RMSE can be either in sample our out of sample. I suggest reviewing these concepts in James et al. *An Introduction to Statistical Learning* (http://www-bcf.usc.edu/~gareth/ISL/book.html, see chapter 2), or Hastie et al. *The Elements of Statistical Learning* (see also chapter 2).
- P 3-9 l 30: It is not the samplers that affect the focus of LURs on long-term averages.
- P 3-10 l 7: I'm not sure how this example makes the intended point. From an epi point of view, better predictions don't necessarily give better health effect estimates. (see Szpiro et al. [2011] *Epidemiology*) So citing an improvement in R^2 doesn't necessarily help us understand whether or not LUR model results will be useful in epi studies.
- P 3-10 paragraph starting line 11: The Atari example discussion could be improved by: saying how the out-of-sample assessment was done, including how many monitors were used for that. Also since

the sampling was for only one 2-week period, how does this reflect other times or a longer averaging period? How would these results be useful in an epi study?

- Pp 3-10 – 3-11: The discussion of Kanaroglou omits mentioning that the data are from a mobile platform and that the spatial autocorrelation was used to improve the model. Further, using the autocorrelation did not improve the predictive ability of the model very much. Also why is there discussion of statistically significant predictors when the focus of the analysis appears to be on prediction of SO_2?
- Section 3.3.2.3: Based on Table 3-1 this section covers kriging too. I see little mention of kriging and the term is not even defined. It should also be clarified whether the authors are referring to simple, ordinary, and/or universal kriging.
- Section 3.3.2.6 (microenvironmental exposure models). Even if researchers publish papers that contradict this point, it is not useful to apply stochastic population exposure models, or any models that simulate exposure, for epidemiologic inference. This is akin to doing multiple imputation without incorporating the outcome variable in the imputation. It leads to biased inference much like the effects of classical exposure measurement error. Stochastic population models are very useful for risk assessment.
- Section 3.3.3 is quite short and is mostly used to reference Table 3-1. Based on the heading, the purpose of this table is to describe metrics relevant only to epidemiology. This is reasonable. (However, this implies that stochastic simulation models should not be included, as is implied in coverage of the micro environmental models.) The new table is a good addition to the ISA, though some work is still needed to refine its content. I suggest that one way to improve this table is to recognize there are two components to exposure assessment for epidemiologic applications: 1) design and collection of exposure data and 2) using the data or deterministic models to produce exposure estimates for study subjects for use of epidemiologic inference, e.g. by using a prediction model. The first column, exposure assignment method, lumps the source of data and the method for coming up with an estimated exposure for an individual in one category. The epidemiologic application column looks useful, but a bit notional. If there is no reference that can be cited to connect to an application being reviewed for SO2 epidemiology (in this or previous SOx ISAs), then I think that particular application should be omitted. The errors and uncertainties column appears to refer both to the exposure metric as well as to the impact on epidemiology. Finally, I suggest that Section 3.3.3 provide more of a road map to understanding each column in Table 3-1 particularly with respect to what criteria are used to fill in the cells of that table.
- Table 3-1: There are some questionable statements made in this table. For instance, that few input data are needed for IDW and kriging (presumably this is ordinary kriging, not universal kriging?) is misleading and if true almost certainly guarantees that the limitation noted (that it doesn't fully capture the spatial variation).
- Section 3.4 (exposure assessment, exposure error, epi inference): In principle this entire section starts off well, but it seems to get bogged down in detail that doesn't necessarily help readers understand the key issues for epidemiologic inference. Perhaps this is my misunderstanding because

instead this section is intended to cover exposure both as it informs epidemiology AND as it informs risk assessment? The intent of this section should be clearer to the reader. Also, in the context of epidemiology, by not distinguishing major study designs early, some sections are more confusing than necessary. For instance, the presence of spatial variation (section 3.4.2.2) makes using a single monitor in a time series study more problematic for SO_2 than other pollutants (e.g. $PM_{2.5}$) because of spatial variability. However, with sufficient data to model it, the spatial variation is an asset for SO_2 in cohort studies. By not bringing up study design before this section, the discussion of spatial variation is less helpful than it would otherwise be.

- Section 3.4.1.1: It would be helpful to provide some overall (perhaps closing) comments to help a reader use the discussion of AER in this section to better understand SO_2 epidemiology.
- P 3-24 line 9: Is this statement true that ambient SO_2 concentrations from central site monitors are commonly used as exposure surrogates in epi studies? I would agree for the time series design, but I need more evidence before I would agree that such a broad statement applies to all designs. I suggest references are needed to the studies reviewed in this document, or a less sweeping statement should replace this one.
- Section 3.4.2.1: A review of activity patterns seems particularly pertinent for risk and exposure assessment. I'd like to see the purpose of this section clarified. Is it for risk assessment, epidemiology, or both? If for epidemiology, what insights do we gain from the material in this section?
- P 3-27 lines 14-16: I suggest the sweeping statement about spatial variation leading to exposure error should be revised. In cohort studies we take advantage of spatial variation in long-term averages to make inference about health effects. Not capturing this well can lead to Berkson(-like) and/or classical(-like) measurement error. But without leveraging this spatial variation we have no information with which to infer SO_2 health effects in cohort studies.
- P 3-27 paragraph starting line 20, as well as much of the remainder of Section 3.4.2.2 on spatial variability: this discussion seems to be referring to time series studies without stating this.
- Section 3.4.2.3. Temporal variability: This is clearly focusing on the time series design. Thus the discussion is easier to follow and better tuned to the goals of the chapter than the previous subsection.
- Section 3.4.3: The Zeger et al. (2000) paper focuses on the time series design. The discussion should make this clear. The properties of the health effect estimate discussed should apply more generally to other designs, so I expect only a modest clarification is needed in this section.
- Section 3.4.3 Co-pollutant relationships: It will be important to understanding to recognize that correlation between measurements at the same location is only one aspect of understanding co-pollutant relationships, at least in time series studies. Pollutants that vary dramatically over space (as SO_2) will have even lower correlation with other pollutants as the distance between locations increases, due to the inherent spatial heterogeneity of the pollutant. I'm not sure whether the full discussion in the subsections reflects this understanding.

- P 3-37 lines 19-20: I would only agree that SO_2 exhibits a relatively high degree of exposure error compared to other criteria pollutants for the time series study setting that relies on one central site monitor, particularly in an area with SO_2 point sources.
- P 3-37 paragraph starting line 24: I assume temporal co-pollutant correlations is referring to the correlation between pairs of pollutants measured at the same location, with repeat measurements over time used to estimate the correlation of these co-located pairs. I am finding it difficult to understand the sentence that starts on line 28: Is this for correlation of spatially varying pollutants, as one might have predictions of two pollutants at various locations in space? And what does "within-pollutant variation across space" mean here?
- P 3-38 lines 4-6: I assume these sentences describe correlations between pairs of measurements **at the same location**. Correct? The captions or footnotes for Figures 3-4 through 3-7 should mention the number of sites included in the analysis. (I recognize this may vary by pollutant, so perhaps this gets added to the side of the figure.)
- P 3-44 line 35: Please specify the within-daily time scale being referred to.
- P 3-45 lines 4-6: I think this sentence is conflating measured relationships with underlying relationships to draw conclusions about epidemiology results. What actually matters is how correlated the population exposure to SO_2 is with other co-pollutants. The observed correlation at a single site may not reflect that, particularly in areas with a large amount of spatial heterogeneity of SO_2 (e.g. due to local sources).
- Figure 3-8: Do I understand correctly that each data point in each boxplot represents the correlation between pollutants as reported in the references listed on page 3-44? And was there any attempt to address whether or not monitors were co-located in this reporting of correlations, at least for the shorter-term measurements? And what do the data in the long-term correlations plot represent? How many studies are reporting in each row? (Do the dots reflect the raw data on these plots? Note: My copy of the ISA does not have any color on Figure 3-8, so it appears there are no red dots shown.)
- Figure 3-9: The data are approximate (because the median year for a study covering multiple years is used) and the scatter is large. I expect none of the regression lines plotted have slopes different from 0. I suggest the Agency consider dropping this analysis, figure, and discussion because it may not be sufficiently informative.
- Page 3-48 and Figure 3-8: I need help understanding the source of the correlations for the long-term estimates. Are these for long-term average predictions for pairs of pollutants at different participant residence locations? Or for co-located measurement pairs across monitors?
- Section 3.4.3.2 would benefit from a statement about the implications for SO_2 health effect studies.
- P 3-49 line 26-7: What is a "surrogate target parameter of interest"?
- P 3-49 line 28: Pure Berkson error (U) is the unobserved part of the true exposure (T). The observed part is X. i.e. T=X+U. Berkson error is independent of X, and has mean 0. See the chapter by Armstrong in Baker et al. (2008): *Environmental epidemiology*.

- P 3-49 line 30-1: Yes, for a linear disease model. There will be some specification bias in a nonlinear disease model. (I suggest just softening the statement to say something more like "Pure Berkson error generally does not bias...." There are some nuances with Berkson-like measurement error.)
- P 3-50 line 2: Replace "but" with "and".
- P 3-50 line 5: Insert "pure" before "Berkson"
- P 3-50 lines 25-27: I think the authors mean that nominal coverage of the CIs is below 95% for exposure effect estimates conditional on mismeasured covariates.
- Section 3.4.4.2 Long-term cohort studies: This section assumes that the ambient concentration value used in these studies is the concentration at the central site monitor. This may be true for older studies, such as the ACS and 6 cities studies, and even the more recent WHI study (Miller et al 2007). However, more recently cohort studies are using predicted exposures from a model, where ambient measurements come from multiple monitors. Further, the theoretical literature on measurement error impacts in cohort studies make the latter assumption that the exposure metric used in the epi study analysis is predicted from a statistical model.
- P 3-56 line 3: For clarity I suggest replacing "estimate" with "prediction"
- Section 3.5 Summary and conclusions:
    - Corrections to previous sections should be brought forward into this section
    - While for discussion of health studies, connection to and changes from the previous ISA seem useful, for the discussion of exposure I suggest this kind of connecting text (line 21 p 3-58 to line 7 p 3-59) is less helpful. I suggest dropping this comparison in favor of better articulating our current understanding of the important features of exposure assessment for application to Chapter 5, and probably also for the upcoming REA and PA documents.

Chapter 5 comments, in particular focusing on the link with Chapter 3

- Pdf page 309, Table 5-9 reference to Sheppard et al. (1999): There was only one SO_2 monitor in this study. (See the map in the paper) It was/is located in an industrial area near a cement plant. So this monitor is not likely to be the best representation of population average exposure. I expect this lack of monitor representativeness impacted the results.
- Overall, in the health outcome summary tables for causal and suggestively causal effects (e.g Tables 5-21, 5-24), it is appropriate to bring in the two exposure-related criteria (uncertainty regarding exposure measurement error and uncertainty regarding co-pollutant confounding). These criteria should have explicit links to the discussion in chapter 3. Furthermore, some of the details in these tables should be reconsidered:
    - Table 5-21, uncertainty regarding exposure measurement error: I wonder whether it would be worthwhile to distinguish studies that are judged to have population-representative monitors for SO2 and those that aren't. Also I think the point is that, for time series studies, the existing fixed site monitors in a given study may not represent population-average exposure.

Time series studies don't address spatial variability well, by design, so the key feature is whether or not the exposure metric used represents the population-average exposure.

- Table 5-21, uncertainty regarding co-pollutant confounding: How much of the distinction in copollutant confounding effects noted could be due to monitor siting? Also, why is the measurement error differential here? Please check the definition and clarify. Finally, if there is differential error "limiting" this inference, this is a topic that should be discussed explicitly in chapter 3.
- Table 5-24, uncertainty regarding potential for measurement error in exposure estimates: This text doesn't inform the reader about what the concerns are. It would be most helpful if the pointer to chapter 3 would be to a section that explicitly discusses the studies alluded to and why, specific to those studies, the concerns noted are problematic.
- Table 5-24, uncertainty regarding potential confounding by copollutants: I wonder whether the table entry is sufficiently informative. What do we know about spatial correlation in the studies cited in the next column? It would be helpful to ensure the table, and the chapter 3 section it references, address the specific issues summarized in this table.

- Related summary table comments on Table 5-35 and Table 5-41
  - Table 5-35 evidence about uncertainty due to confounding by correlated pollutants: The evidence cited seems reasonable as an explanation for why this factor contributes to a lack of causality. It would be helpful to have a chapter 3 reference in the next column.
  - Table 5-35 uncertainty due to exposure measurement error: This category does reference a few specific studies, which is helpful.
  - Table 5-41 uncertainty regarding potential confounding by co-pollutants: The second sentence is a reasonable summary that might be carried into other short-term studies summary tables.
  - Table 5-41 uncertainty regarding exposure measurement error: The key idea is that SO2 is heterogeneous over space. It is less important how monitors are correlated, though clearly the observed correlations reflect the underlying spatial heterogeneity. As I noted above, I believe the issue is whether the monitors used in the time series studies are representative of the population-average exposure.
- Chapter 5 topics that perhaps also belong in chapter 3
  - Measurement error may affect the shape of concentration-response function estimates
  - Averaging time matters and the actual averaging time used may not be aligned between the exposure assessment (i.e. development of the exposure metric to be used in the epi study) and the inferential goals epi study.

Dr. Frank Speizer

Executive Summary and Chapter 1

The level of detail and formulation is excellent. In particular, the description of the what constitutes the standard seems better laid out than my memory of previous documents. In addition, the logic of how EPA has got to a 5-minute average is well justified.

Page xliv, line 15-18: This is a confusing sentence in that the first half talks of levels outside the US and compare them to in the US.

Page xlvii, line 7-8: suggest editorial modification (**see bold)** … the **potential** mode**s** of action underlying these responses **are** uncertain.

Table ES 1: This is an effective summary of key findings and by providing specific references to table is subsequent chapters allows for ease of documentation.

Page 1-27, Sentence on lines, 8-10. Although the statement is probably accurate, I am not sure it can be fully justified from the data presented in paragraphs above. The issue is that because of the very short term effects of acute exposures, it is really the case that the epidemiologic data may simply not be applicable to judging the shape of the dose response curve since the minimal time measure in most of the epi data is 8-24 hours and the acute response may be short lived over 5-10 minutes and may with repeated exposure be either potentiated in some individuals or diminished due to not being responsive to repeated exposures. (Will need to check in Chapter 4 for evidence of both).

Section 1.8 Summary: Succinct and well written and understandable to lay public

Chapter 4

Page 4.3, Table 4.1. Need to fix age column. Original source has data to above 61.

Chapter 5

General Comment: Overall this chapter had done an excellent job of identifying the levels of causation associated with the major potentially significant outcomes as related to both short and long term exposure. Where these have remained consistent with the 2008 ISA and where they have changed, both in response to the literature as well as our suggestions from the last review, the documentation is well presented. There remains an issue that was discussed at the meeting related to the fact that the data obtained in asthmatics may not be generalizable to all asthmatics as most the studies have been done in

mild young adults. Thus the findings are on the conservative (more optimistic) side and this must be taken into account when considering what the effects might be in more severe, older or disadvantaged groups or groups with greater susceptibility. For example, we know that the underlying rates of asthma in children are greater in African Americans and selected Hispanic groups compared to whites. (Perhaps this latter comment is more for the RA than the ISA.

Specific Comments

There seems to be some inconsistency in the supplementary tables. Table 5S-1 Summary of epidemiologic studies of SO2 exposure and other morbidity effects (i.e., sensory, nervous and gastrointestinal and other effects (3420016) provides descriptive outcomes; however Table 5S-1 Summary of epidemiologic studies of SO2 exposure and other morbidity effects (i.e., eye irritation, effects on the nervous and gastrointestinal systems).(3001861) which I believe is the same studies is providing risk scores with outcomes mentioned in the description of the study. This seems to be unnecessary redundancy and in fact by labeling each table as 5-S1 is confusing.

Section 5.1.2.1, line 16: suggest add to list of disciplines human clinical studies (since a significant part of the causal inference is from these studies).

Although I understand why Table 5-1 is included I do not find it very useful. Most of the actual outcomes are in tables 5-2-5-4, and the text in between. In addition, most of the studied pre-date the 2008 report, which makes Table 5-1 redundant.

Page 5-23, lines 31-32: Not clear that the assumption that methocholine response is equivalent to SO2 exposure. SO2 is more likely to be a vagal response whereas methocholine is more directly smooth muscle stimulated response. Suggest sentence be made more circumspect. In fact, suggest the whole paragraph be considered for redrafting as it seems to mix a number of potentially quite different mechanisms (e.g. allergy, obesity, mouth breathing, boys> girls, age).

Page 5-25, line 1-2. The concluding sentence of the paragraph does not follow from the data presented. As presented and in the articles cannot conclude that SO2 cause the increased responsiveness to dust mites allergen. Footnote is redundant.

Page 5-28, Table 5-5 Make post 2008 part of title of table.

Page 5-30, line 1-3 “which may better represent some component of exposure than a monitor not sited in a subject’s microenvironment.” Editorial comment not necessary

Page 5-30 paragraph lines 17-36. Suggest this paragraph be reconsidered and re-written. It seems to be a random cataloguing of findings without much logical thought. It mixes space, time, lag, diurnal variation, atopy and allergy in animals and human and draws a sweeping conclusion that is not helpful.

Table 5-6 Ditto title change as suggest to Table 5-5 Far too much detail to be useful. Last Column could be a yes/no or what the co-pollutant models show for SO2 rather than details.

Page 5-44 line 12: I can find no documentation for the 90 ppb in text of any of the papers cited. In fact according to Table 5-8 in several of the recent studies that were in fact about auto traffic pollutants and diesel, no SO2 measures are actually recorded and it looks as if estimates for SO2 are made on the basis of correlation data with EC as discussed on page 5-45. Please check.

line 32: Please check if you mean ”less uncertainty” or “more uncertainty”

Page 5-47 line 7: I don’t think the word “uncertain” is appropriate here. Clearly there is potential confounding and potential interaction; and thus the relation to SO2 is complex but not uncertain (the way the EPA generally defines uncertainty).

Line 31: Ditto. Again uncertainty is equated with increase variability and the terms mean different things.

Page 5-62 Para beginning line 24 (and elsewhere where the Zheng [2015] article is referred to) my reading of the article is that the authors found the potential for interactions between pollutants too complex to consider analyzing and thus only presented single pollutant models. I think they did it right and their interpretations are correct but from the standpoint of indicating what they did this needs to be indicated in some way to avoid unnecessary criticism.

Page 5-67, para beginning line 21: Not mentioned in the paragraph but apparent in Figure 5-5, is the remarkable consistency of the SO2 effect across all models. I think this should be stated.

Page 5-71 at end of section, it may be worth mentioning that the effects seem to be more consistent during the summer vs winter. This raise a couple of issue: 1) kids may be outdoors more during the summer with higher ventilator rates and thus greater exposure; 2) however, since SO2 levels are stationary source pollutants and depending on region might be considerably higher In winter than in summer, but this may be true for some of the other pollutants, thus the finding that the SO2 effect remained constant over different multipollutant models is somewhat reassuring. (don’t feel strongly that this has to be included).

Page 5-74, line 12: “imprecise” is probably not the right word. The effect on SO2 on TBAR is not significant after adjusting for PM2.5. In fact, because of the relatively high correlation between the two (in contrast to the other gases) one cannot assess which is acting (since the sum of both is about the same as SO2 alone.

Page 5-82, line 20-21: Again the word “imprecise” I understand that the author is looking for a word to indicate that the Confidence Interval crosses 1.0, but I am not sure this is the best way to convey it. In this case [−0.82% (95% CI: −1.9, 0.31) per 10-ppb 21 increase in 2-h avg SO2] the suggestion is that the effect is really null or at best a modest negative association that is not significant. I don’t think the word “imprecise” conveys that. Note in other places (e.g. page 87, the words used for similar statistics are

"limited evidence," page 89: "large 26 uncertainty estimates") This may be an issue that the reviewers should discuss.

Page 5-85-86. Minor point: Suggest reverse order of Figure 5-7 and Table 5-12, which would allow reader to see studies and concentrations before risks.

Page 5-91, line 9-11. I take some issue with the first half of this statement. The fact that all the studies do not point in the same direction might lead one to say "inconsistent." However, taken together one could argue that there is reasonable biologic and epidemiologic plausibility of an increase risk. I would agree with second half of the effects of attenuation by PM.

Page 5-99 Summary of Respiratory Infections. Somewhere in this section (perhaps later) something needs to be said about the comparison between respiratory infection in the developed world vs the developing world and in places with different health care delivery systems. The risk of infection causing ED visits and Hospitalizations may be quite different and as is noted in Figure 5-8 where the Canadian Cities studies have considerably lower risk that other places. (This is mentioned on page 5-106 in discussion of South Korea).

Page 5-101 Figure 5-9: Some reordering or at least change in symbols should be made to separate out children from adults over 65.

Page 5-127, and Table 5-19: I am afraid fall back into the trap of having the need to report everything. There is nothing in either the table or the text in this section that even approaches the need to have been included. This leads to frustration for the reader as these studies probably could have been left out and not reported. At this stage can leave in but It is this kind of reporting that I thought the ISA was to get away from.

Page 5-134, sentence begins line 10: **This seems inappropriate at this point.** Much of the data previously presented has been from Asian countries and to say now that the mortality data specifically should be looked at with caution because it is Asian call into question much of the previously reported work. Suggest let the mortality data stand without this caveat.

Page 5-135 Figure 5-10: Add to title of Figure "in 4 Chinese Cities"

Table 5-21: This is a very effective summary, congrats!

Page 5-145-6 and Table 5-22: The text indicates the effect estimate for the Nishimura study is per 5ppb change in annual average. If so this need to be added to table text as it currently indicate the effect size with specifying pollutant change. Assume Clark study is the same. For several of the other studies cited, much of the detail under the selected effects column belongs elsewhere (for example, the whole last paragraph for the Tam study could go under the first column.)

Page 5-150, line 15-16. Not clear why this sentence is included since the Borrell study makes no mention of pollution and does not suggest that either obesity is a confounder or is interacting in the pollution asthma pathway.

Need to keep this in mind. P 5-156 Thus, multiple lines of evidence suggest that long-term SO2 exposure results in a coherent and biologically plausible sequence of events that culminates in the development of asthma, especially allergic asthma, in children.

Page 5-158 mention of Table 5S-12. It is not clear why some of the supplementary tables are in as S. For example, this particular table I would have thought belongs in full text. (I may have missed a statement that defines why tables are where they are.

Page 5-159. The recent studies support conclusions of <u>no association</u> between long-term SO2 exposure and lung function in children made in the 2008 SOX ISA (U.S. EPA, 2008d). Not sure I agree and would like to discuss.

Page 5-176 Overall, despite some epidemiologic evidence of an association between short-term exposure to SO2 and hospital admissions and ED visits for ischemic heart disease and MI, uncertainties regarding copollutant confounding continue to impede the determination of an independent SO2 effect.

Page 5-182, Figure 5-13: I read this figure as showing only one study with significant results. (Vancouver study) and looking at the original paper there are lots of analyses with only one or two significant results. In addition, results seem to be confined to females only. Suggest it be played down even more (Page 5-179. Line 31).

Page 5-184 As such, the current evidence does not support the presence of an association between ambient SO2 and blood pressure.

Page 5-185 Given the limited epidemiologic evidence, the association between ambient SO2 concentrations and venous thromboembolism is unclear

In summary, the available epidemiologic evidence is limited and inconsistent, and 6 therefore does not support the presence of an association between ambient SO2 concentrations and hospital admissions or ED visits for heart failure.

Page 5-199 Lines 4-10 reproduced below. Not clear what sentence in bold from lines 8-10 means Limited analyses of model specification, the lag structure of associations, and the C-R relationship suggest that: (1) associations remain robust when alternating the df used to control for seasonality; (2) associations are larger and more precise within the first few days after exposure in the range of 0 and 1 days; and (3) there is a linear, no threshold C-R relationship, respectively**.** However, for both total and cause-specific mortality, the overall assessment of linearity in the C-R relationship is based on a very limited exploration of alternatives.

Page 5-200-1. Discussion of HRV. Need input from physiologist. HRV change may be more important than whether it is positive or negative and if so discussion on these pages should change to be similar to what follows on experimental studies

Page 5-202. Overall, studies evaluating the effect of ambient SO2 concentrations and 14 measures of HRV and heart rate remain limited.

The two reviewed studies provide limited evidence of association between short-term SO2 exposure and markers of ventricular repolarization

Page 5-207-8. The experimental data although conducted as significantly (logs higher) higher levels are impressive in trying to understand mechanisms, and are surprisingly followed with a summary statement that is not consistent with these findings. Surely whoever wrote the conclusion was thinking of the human data and that is understandable but it is not complete and is ignoring the data presented in paragraph above.

Page 5-208 section: **5.3.1.11 Summary and Causal Determination** Continuation of above. The first sentence is justified but not because of no biologic plausibility. Clearly most of the data are inconsistent and not adequately adjusted for co-pollutants and measurement error in exposure. However, the biologic plausibility is impressive and cannot be used as an excuse. Suggest re-write. Ditto page 5-213, line 5-8.Going on the use of the term “lack of coherence” between human and experimental studies seems too strong. If in the author believes because the HRV changes go in different directions that is not enough. If he/she thinks that the animal data is irrelevant because of exposure level say so but the data are significant. We will need to discuss.

Would have made tables 5-32 and 5-33 Appendix tables

Table 5-34, could have combined the two Dong studies since almost all the words are the same and results could have been listed.

Page 5-229. In conclusion, the evidence lacks coherence and is of insufficient consistency, and thus, is inadequate to infer the presence or absence of a causal relationship between long-term exposure to SO2 and cardiovascular health effects.

Page 5-232, line 1-5: This sentence seems to have no place in this document. Suggesting a report may be coming makes little sense.

Table 5-36: The last column in this table has risk rates that have been considered for other outcomes to be important. What is missing for the most part is the fact that for these rates there is no adjustment for co-pollutants. Suggest either add this to last column or add another column for Comments in which whether multipollutants were or were not considered.

Page 5-243, lines 14-15. This summary statement is too strong as most of the data presented are null or non-significant.

Page 5-247 For consistency sake section on Infant Mortality needs a summary statement.

Starting page 5-259, Table 5-39. There are 3 entries that do not give levels of exposure. At least for the Moogavkar one would have thought the range would be the same as for Dominicii as the same data base is being used. I could not get to the Bellini or Atkinson papers, to details of the aerometrics. I find it hard to believe the MISA2 paper did not have them for SO2. In any case without some estimate of exposure hard to understand how the calculations for these papers is made in Figure 5-17 and 5-18.

Page 5-363, Table 5-40. Please clarify if these are two separate 2 pollutant models or NO2 is added to model with SO2 and PM. The latter would make more sense, and the former would suggest that the confounding is extreme, and suggests that SO2 has no effect.

Page 5-278 In conclusion, the consistent positive associations observed across various multicity studies is limited by the uncertainty due to whether SO2 is independently associated with total mortality, the representativeness of monitors and the 24-h avg SO2 exposure metric in capturing the spatial and temporal variability in exposure to SO2 (Section 3.4.2.2 and Section 3.4.2.3), and the uncertainty in the biological mechanism that could lead to SO2-induced mortality (Section 4.3). Collectively, this body of evidence is suggestive, but not sufficient to conclude there is a causal relationship between short-term SO2 exposure and total mortality.

Page 5-293 The overall evidence is inadequate to infer a causal relationship between long-term exposure to SO2 and total mortality among adults.

Page 5-304 The overall evidence for long-term SO2 exposure and cancer is inadequate to infer a causal relationship.

Chapter 6

Page 6-3--4 Pre-exisiting Disease/Condition, Table 6-2. I am also a bit confused by the table. I assume the N is the total population. (234,921 +6,292= 241,213). Is this total pop of US in 2012? Secondly, concern here that the description is incomplete as it describes Asthma but doesn’t say anything about COPD, which in the older age groups amount to about 3.5 million people. Further with regard to potential significant risk this group may have a greater impact on health care/delivery/utilization system than the larger asthma group

Page 6-4, line 23: Suggest change age range from 18-20 to 18-25, as male continue to grow past females.

Page 6-5-6, Conclusion line 32-33 and Table 6-4; Need to indicate in table that XX/d represents counts per day of ED visits? Or something else if not. In fact, the table is not really interpretable and not clear what is meant by conclusion as in some cases both cases and references are children, in other they are not and if these are counts per day and the comparison is between younger and older children **on the same days** the differences really related to the population base of the number of children in catchment area rather than modification by SO2. Please clarify.

Page 6-9, Table 6-5. Ditto same problem.

Page 6-16. In conclusion, evidence is adequate to conclude that people with asthma are at increased risk for SO2-related health effects. Asthma prevalence in the U.S. is approximately 8−11% across age groups (Blackwell et al., 2014; Bloom et al., 2013), and thus, represents a substantial fraction of the population that may be at risk for respiratory effects related to ambient SO2 concentrations.

My problem with the conclusion is that no real estimate of population at risk is made for any of the potential risk groups and thus much of the chapter is really a rehash of data in Chapter 5.

Dr. James Ultman

Executive Summary and Chapter 1

The executive summary has been shortened by removing redundant material, and has been made more accessible to the non-technical reader.

Chapter 4

Several revisions/additions have led to an improvement to this chapter. Entirely new sections on the structure/function of the respiratory system and breathing rates/habits provide a improved foundation for the later sections on SO2 absorption and possible mode of action. The inclusion of material on the possible effects of obesity on SO2 absorption vis-à-vis modification of breathing habit is also recognized.

1) In the section on chemistry, the term "Henry's law constant" for SO2 (pg 4-8, line 4) represents the ratio of molar SO2 concentration in air to the equilibrium SO2 concentration in water, which appears as dissolved gas and as reversible reaction products (Eq. 4-1). Strictly speaking, Henry's law constant does not include reaction products. Thus, instead of using "Henry's law constant" for SO2, the authors are urged to define an "effective Henry's constant", as was done in the original Tsujino (2005) article. For ozone, which does not undergo a reversible reaction in water, the value given in the text is a true Henry's law constant.

2) In the section on absorption, the SO2 mass transfer rates given on page 4-9 (lines 3-5) comparing infants and young adults seem to be based on computations made by the authors of the ISA. The computations rely on an equation for mass transfer rate (Asgharian, Eq. 3) that requires the concentration of SO2 to be known at the respired gas-ELF liquid interface. It appears that this concentration has been neglected so that local absorption is proportional to the gas-phase Sherwood number=(airway diameter)(gas phase mass transfer coefficient)/(gas phase diffusion coefficient). This is not necessarily the case; it is more likely that a transport resistance modulated by diffusion-reaction processes in the ELF result in a non-zero interfacial concentration which opposes absorption. If the current analysis of uptake rates is retained in the revised document, a justification should be provided for neglecting SO2 interfacial concentration. By the way, the assumption of a zero interfacial concentration is not consistent with the occurrence of SO_2 desorption during expiration, which is asserted in other places in the chapter.

3) Also in the section on absorption, the statement "dose as ventilation per bronchial surface area" (pg. 4-10, lines 3-6) is vague. I don't think that the authors are referring to actual dose per unit surface since this will depend on additional factors as well as ventilation. More specifically, ventilation can be thought of as a surrogate for inhaled dose (i.e. ventilation×inhaled concentration). Thus, in the context of comparing two individuals of different ages exposed to the same inhaled SO_2 concentration, "ventilation

per unit bronchial surface" represents their relative "inhaled doses per unit bronchial surface."

4) As a framework to address comments 2 and 3, the revised ISA should include a more comprehensive conceptual description of how transport processes transform "inhaled dose" at the airway opening into "uptake" into a local target tissue. This transformation involves longitudinal convection and diffusion processes in the respired gas phase as well as lateral diffusion and reaction processes in the underlying ELF. In a simple model, local uptake is proportional to the difference in local pollutant concentration between the gas and tissue phases. The proportionality constant is an overall mass transfer coefficient that depends on gas-phase mass transfer coefficient, physical solubility, liquid-phase molecular diffusion coefficient, liquid phase reaction rate coefficient and ELF layer thickness (for example, see page 268 in Hu, S.C.,et al., 1992, Comput. Biomed. Res. 25: 264-278).

Dr. Ronald Wyzga

Charge # 5 - Populations and Lifestages Potentially at Increased Risk for Health Effects Related to Sulfur Dioxide Exposure

Please comment on the adequacy of these revisions to clarify the characterization of the evidence for increased risk of S02-induced health effects in different populations and lifestages.

I am a bit disappointed by this chapter: first of all, it needs to clearly state what all of its objectives are and how its contents/conclusions will be used; secondly, it mimics much of the information in the preceding chapter without really adding any new perspective; finally, it could provide more detail that would help define all of the conditions for which health risks are elevated.

Section 6.3.1

In discussing asthmatics, it is important to identify those behavioral, environmental, and physical characteristics that could exacerbate asthmatic response, such as the presence of exercise, not being medicated, cold weather, or being obese. This is not to minimize the possibility of asthmatic response, but it could provide information both to asthmatics and to the public health community about those conditions when as adverse response is more likely.

Section 6.5.1.1

One reason that children may be more susceptible is that they spend more time outdoors and that they exercise more frequently.

Section 6.5.3 Since ambient levels of SO_2 are tied to specific point sources, those with lower socio-economic status may live nearer to these sources as neighborhoods near sources may be less desirable. In addition, some subpopulations, such as black children have a much higher prevalence rate for asthma than other subpopulations.

Other Comments:

Executive Summary (and Chapters 1 and 2): It is noted that emissions have decreased considerably from 1990 to 2011 and that concentrations of the annual 99th percentile have decreased noticeably from 2011 to 2015. When will emissions estimated beyond 2011 become available? It would be of interest to note that they have also decreased in the most recent period. I note the Dr. Chow's comments present a more recent estimate of SO2 emissions; as a minimum these should be incorporated into the document.

Chapter 1:

p.1.8, l. 14: insert “parts of” before “the West Coast”

p. 1-9, l. 5: Something should be said about the performance of these models here.

p. 1-10, ll. 1-7: what about the relative concentrations between ambient and indoor levels? This as important as the correlations.
l. 12: What is “moderately correlated?”

p. 1-12:ll. 12-14: This sentence confuses me. Why are they “most informative” when measured levels are available?

p. 1-17, l. 8.: What is meant by “moderate decrement?”

p. 1-27, ll. 19-30: Measurement error can also complicate/potentially bias estimates of the shape of the dose=response curve. Since this is referred to later, it should be mentioned here. Another contributor to measurement error is the possible incorrect measure of exposure in epidemiological studies. Human clinical studies demonstrate changes among asthmatics after exposures as short as 5 minutes. Yet most epidemiological studies consider a 24-hour average of SO2 (or possibly the maximum 5-minute concentration); hence if a period of exposure less than 24 hours is relevant, the use of a 24-hour average is incorrect and subject to measurement error.
p. 1-29, ll 24-28: Children may also be at increased risk because they spend more time outdoors and exercise more often.

Chapter 2:

p. 2-1: Are there any data available to update Figure 2-1? Concentrations have declined from 2011 t0 2015.

p. 2-74, l. 4: Delete “good” as it is subjective and within a factor of two may not be “good” in the minds of some readers.

Chapter 5:

p. 5-17, ll, 26-27: Are these cutoffs defined to be the level of adversity?

P, 5-30, l3.: Here the co-pollutant issue could be more important as on-road sources, including SO2 from diesel emissions could be more highly correlated.

p. 5-34, l. 19: What is an “imprecise association?”

p. 5-35, section titled Respiratory Symptoms in Populations with Asthma: There should be some attempt to couple the symptom results with the lung function results

p. 5-39, l.4; symptom "categories?"

p. 5-63, ll. 1-8: This result could also be due to the fact the individuals may spend more time outdoors and exercising in the summer (often vacation) months.

p. 5-65, section titled Concentration-Response Relationship: The fact that measurement error can influence the estimated shape of a dose-response curve need be stated. See also comments for p. 1-27.

p. 5-71, ll. 31-36: See above comment.

p. 5-109, ll. 10-15: There could also be behavioral differences among locations as well; e.g., amount of time outdoers, exercise levels and frequency, use of air conditioning, etc.

p. 5-135, ll. 1-9: See comment for p. 5-65.

p. 5-144: Mention is made of the several positive studies; while statistical significance is not the be-all and end-all, it would also be helpful to learn how many of these studies showed significant results.

p. 5-150, ll. 22-23: See above comment.

p.5-261, l. 10: See above comment.

p. 5-263, l. 11: See above.

p. 5-271-ll. 13-26: See comment for p. 5-65.

p. 5-284, ll 5-6: What does "positive, yet imprecise" mean? Positive but not significant?
l. 9-14: I would worry about EC and VOCs as possible confounders in this study.