UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON D.C. 20460

OFFICE OF THE ADMINISTRATOR
SCIENCE ADVISORY BOARD

January 9, 2008

EPA-CASAC-08-005

Honorable Stephen L. Johnson
Administrator
U.S. Environmental Protection Agency
1200 Pennsylvania Avenue, NW
Washington, DC 20460

Subject: Clean Air Scientific Advisory Committee's (CASAC) Peer Review of EPA's Integrated Science Assessment (ISA) for Sulfur Oxides – Health Criteria (First External Review Draft, September 2007)

Dear Administrator Johnson:

The Clean Air Scientific Advisory Committee (CASAC or Committee), augmented by subject-matter-experts to form the CASAC Sulfur Oxides Primary NAAQS Review Panel (hereafter referred to as the panel) conducted its review of EPA's *Integrated Science Assessment (ISA) for Sulfur Oxides – Health Criteria* (First External Review Draft, September 2007) on December 5 - 6, 2007. The ISA for sulfur oxides (SO_x) was produced for EPA's new process for reviewing and revising National Ambient Air Quality Standards (NAAQS). Done properly, the ISA should be an informative, succinct, and useful summary of the evidence for consideration of the NAAQS.

Overall, the panel found that the first draft ISA for SOx is well developed, and with further revisions, should address CASAC's concerns. We recognize that an administrative decision has been made to consider the sulfate particulate matter as part of the PM NAAQS review process; therefore, sulfates will only be considered here to the degree that they inform the consideration of the potential health effects of sulfur dioxide (SO_2). Major points from CASAC panel members are summarized below. Individual recommendations from CASAC Panel members to strengthen the next draft are appended in Attachment B.

1. To what extent are the atmospheric chemistry and air quality characterizations clearly communicated, appropriately characterized, and relevant to the review of the primary SO_2 NAAQS?

Some improvements need to be made to the discussion of the atmospheric chemistry of SOx. For example, the low saturation vapor pressure of sulfuric acid must be mentioned as the driving force for the gas-to-particle conversion of sulfate. Also, a more comprehensive discussion of the oxidation of SO_2 to sulfate in coal-fired power plant plumes is needed since this source category is currently the major emission source of SO_2 in the United States.

Ship emissions are a major issue near ports and along shipping lanes and as sulfur emissions from point sources and on-land mobile sources decrease, ship emissions take more relative importance. Accordingly, it would be useful to discuss ship emissions (and other relevant SO_2 emission sources), including associated uncertainties.

2. Are the properties of ambient sulfur oxides appropriately characterized, including policy-relevant background, spatial and temporal patterns, and relationships between ambient sulfur oxides and human exposure?

There is a need for a better description/characterization of the urban spatial heterogeneity of SO_2 concentrations, including contrast of urban/rural concentrations and characterization of concentrations with respect to current monitor locations (e.g., what is the average concentration at source-oriented monitors versus at population-oriented monitors and at rural background monitors). Quantification of spatial heterogeneity and how this impacts our ability to assess area-wide ambient concentration patterns, including five minute peaks, is also needed.

The lack of correlation between SO_2 and sulfate is not surprising considering the slow conversion of SO_2 to sulfate on average. Correlations should be presented between SO_2 and other air pollutants such as NO_2, $PM_{2.5}$ and $PM_{10\text{-}2.5}$, and specific components of particulate matter (e.g., elemental carbon and specific metals).

Additional information is needed to compare the distributions of monitoring data averaged over the different times considered for future standards (e.g., 5 minute, 15 minute, 1 hour, 1 day and 1 year). It would be useful, for example, to correlate 5 or 15 minute averages and maxima with the 1-hour averages.

3. Is the information provided on atmospheric sciences and exposure sufficient for the evaluation of human health effects of sulfur oxides in the ISA?

A thorough analysis of the Air Quality System (AQS) and other data should be added to the Annex. A better understanding of the spatial distribution of population exposure is needed. The AQS data characterization needs to more directly address the spatio-temporal variation in SO_2 and the contribution of monitor location characteristics to the summary data because this has direct bearing on the interpretation of the epidemiological studies. There is a paucity of ambient monitors and they may represent near-source impacts rather than average population exposure. For example, how can we estimate population exposure in areas where there is no SO_2 monitor or where the SO_2 monitors are located to characterize the potential impacts of SO_2 sources? Land-use characteristics (e.g., industrial, urban, rural) may help address this issue. Relevant features should be brought forward in Chapter 2.

The draft ISA needs to better address how accurately we can predict personal and population exposure using ambient monitoring. In addition to understanding the spatial distribution of population exposure, a better understanding of the relationships between outdoor, indoor and personal exposure is needed; it will be a useful background for the subsequent use of APEX, EPA's exposure assessment model, for population exposure calculations.

A discussion of the bias and uncertainty associated with the SO_2 monitoring network is needed because of the implications for the assessment of population exposure to SO_2. In particular, the error is likely to be large for population exposure to high SO_2 concentrations.

CASAC asks for a description of the database that exists for 5-minute peak concentrations. A figure should be added to depict an example of the probability distribution of 5-minute average SO_2 concentrations. Plots of similar form (e.g., cumulative distributions) should be provided for annual, 24-hour average, 1-hour average and 5-minute average of SO_2 concentrations. Correlations between peak 5-minute averages, peak 1-hour averages and 24-hour average concentrations should be provided. These correlations could facilitate understanding of the relationships between findings of human clinical studies and epidemiological studies.

A deeper discussion of the physical and chemical interactions between co-pollutants with SO_2 would be useful. In particular, interactions of SO_2 with particles should be discussed.

4. To what extent are the discussion and integration of evidence on the health effects of sulfur oxides from the animal toxicological, human clinical, and epidemiological studies, technically sound, appropriately balanced, and clearly communicated?

The presentation of the results of the animal toxicological, controlled human exposure, and epidemiological studies that have been reviewed is generally technically sound. However, the criteria for selection of specific studies presented in each of the three categories should be clearly stated. In addition, the criteria for judging the strength of findings from specific studies, determining consistency of results, and assessment of aggregate findings of studies on relevant research questions, should also be clearly stated. The existence of publication bias and its consequences, both positive and negative, should be assessed. It should be noted that several relevant studies have not been included in the draft; some of them are recent publications and others did not emphasize SO_2 in their conclusions. These studies are listed in individual comments. The assessment of the aggregate findings needs to be reviewed in light of the approach to the evaluation of evidence of causality to be described in a revised ISA.

The epidemiological data are relatively consistent and coherent with regard to the association of short-term exposure to SO_2 and emergency department visits/hospitalizations for asthma and all respiratory diseases, particularly among children. There is not sufficient discussion of the fact that a substantial fraction (15-30%) of the population may be driving the effects noted in non-asthmatic subjects. This requires additional discussion, since null results across larger (non-asthmatic) population groups may obscure the identification of actual susceptible sub-groups within. Additionally, positive results may be driven by this susceptible

non-asthmatic sub-group, possibly resulting in over-interpretation and assignment of a general population effect.

Interpretation of all epidemiological studies of SO_2 effects (both long-term and short-term) needs to include the context of exposure assessment, particularly with respect to local source monitor siting and how evenly distributed SO_2 is in space. There needs to be a standardized protocol for summarizing relevant studies for the ISA and in selecting which information to report. Characterization of epidemiological study results from multi-pollutant models (such as those involving analytical partitioning between SO_2, PM, and other pollutant effects) should be re-visited to clarify the current presentation and interpretation. Thoughtfully interpreted discussions in the context of (a) interactive effects and effect modification; (b) mediated effects; (c) confounding possibilities; and (d) independent effects need to be expanded in the analysis of epidemiological studies involving two-pollutant and multi-pollutant models.

The controlled human exposure data presented in the current version of the health effects chapter indicate that asthmatic individuals are especially sensitive to SO_2 exposure in terms of respiratory symptoms and bronchoconstriction. While these data do provide some plausibility for the epidemiological studies reporting associations between ambient SO_2 and emergency department visits or total hospitalizations for asthma, they are not informative as to how SO_2 exposure might induce respiratory hospitalizations in non-asthmatic individuals. Plausible mechanisms for other respiratory associations should be considered as well.

More attention needs to be given to the results from the older human clinical studies to leverage new insights from previous studies. This approach may be informative as to who is responsive and under what conditions (levels, length of exposure, co-exposures, severity of health status, etc.) that responsiveness is present. This information needs to be summarized in the document and integrated with the epidemiological results. In addition, recently published additional studies need to be identified and included, as appropriate, in the review.

There are places in the text that need to be tighter, less redundant, and more thematically organized (i.e., each section should have a story line). In particular, the summary/integration subsections should provide an overview of the quantity and quality of the evidence for the health outcome(s) of interest as well as evaluation of how well the toxicological and clinical data support the epidemiologic findings. Although there is substantial information for a number of different health outcomes, the focus of the integrated summary of the findings, the strength of those findings, and confidence in the findings are not clearly presented. A summary table listing the various outcomes, a determination of the level of confidence in the findings, and the implications of the findings would help guide the reader to a cumulative sense of the current aggregate status of SO_2 knowledge.

Much of the material in Chapters 4 and 5 might be brought forward to Chapter 3 to better inform the discussion on susceptibility and vulnerability (two separate issues currently presented in multiple chapters that also should be collected and discussed in one chapter location).

A section of the document needs to be added to summarize current levels of uncertainty and to discuss key areas of research needed to reduce those uncertainty levels.

5. To what extent does the integration of health evidence focus on the most policy-relevant studies or health findings?

The document does a commendable job of selecting out the key documents since the last review, but EPA staff and consultants should more carefully re-evaluate whether there are any key studies prior to the last review that provide especially important insights, or that can be reconsidered in the light of new evidence, in order to aid in the present evaluation of the health effects of SOx. This was especially noted as being important in the review of the toxicological and human clinical studies done on sulfur oxides and co-pollutants.

6. What are the views of the Panel on the conclusions drawn in the draft ISA regarding the strength, consistency, coherence and plausibility of health effects of sulfur oxides?

The EPA staff is to be commended for producing a summary of findings on SOx that is a significant improvement over the same section of the NOx document. Chapter 5 properly summarizes the conclusions from earlier chapters, and the conclusions that are presented are for the most part clearly relevant to the various risk assessment and policy questions that will arise in later activities related to establishing a NAAQS for SOx. Still, there is room for improvement in several areas:

- Chapter 1 contains a series of six policy questions that the ISA is intended to inform. It would be useful to structure Chapter 5 to specifically address these six questions, and then to phrase the final concluding statement - about whether effects are occurring at existing ambient levels – as a natural consequence of this structured assessment of the information. At present, it is not clear how the concluding statement is related to the other findings in the chapter.

- The hierarchy of causal claims used in Chapter 5 is appropriate, but the criteria used to satisfy each of the categories of causal strength are not well specified and in some cases do not comport with best scientific practice. This aspect of the chapter can be improved, especially with respect to criteria of coherence of evidence and robustness of conclusions. A complete description of the approach to causal inference should be provided in a revised ISA.

- The conclusion that evidence of adverse effects at ambient exposures is robust (p. 5-16) appears to be too strong a statement, given the evidence presented in this draft. We recommend a slightly less emphatic conclusion given problems associated with biological plausibility in the face of heterogeneous information from the various categories of data (epidemiological, clinical and animal toxicological). The strongest conclusions should be reserved for those responses with the greatest coherence across the available evidence.

- The chapter should better summarize the limitations in conclusions about causal connections at ambient levels that are introduced by problems of confounding and

weakness in exposure assessment underlying epidemiological studies. In general, more attention should be given to how the strengths and limitations of the existing information, from concentration to exposure to effects, affect the ability to form definitive answers to the risk and policy questions in Chapter 1.

With these improvements, Chapter 5 should provide a good template for the way in which this chapter should be written for future NAAQS ISAs.

7. **What are the views of the Panel on the appropriateness of public health impact and the characterization of groups likely to be susceptible or vulnerable to sulfur oxides?**

This is a good first draft that has correctly identified and discussed data showing which respiratory effects (hospital admissions, ER visits, and acute bronchoconstriction) are the adverse health effects of concern for SO_2. Because of the overlap of Chapter 4 with Chapter 3, it is recommended that the discussions of susceptible subpopulations be moved to Chapter 3 and that Chapter 4 focus on discussions of concentration-response relationships, vulnerable populations, and the potential size of the population at risk. The nature of the dose-response relationship of SO_2 concentration with adverse health effects at current ambient levels is critical to the estimation of the burden of disease imposed on the population. Chapter 4 addresses dose-response relationships as addressed in individual studies, both the human clinical studies and selected epidemiological studies. Toxicological findings that might be relevant are not integrated into this discussion although the majority of animal toxicology data are at SO_2 concentrations that are much higher than ambient levels. Additionally, Chapter 4 should more clearly address the general difficulties of exploring the existence of thresholds and the form of the dose-response relationship at ambient levels and possible influences of co-pollutants in epidemiological studies. Uncertainty issues and variable sensitivity in the populations (i.e., presence of responders) might be addressed more systematically as well.

The discussion of the extrapolation of the clinical studies, involving generally healthy or asthmatic volunteers, to ambient exposure levels needs expansion with a deepening of the discussion of underlying mechanisms and their potential implications for the dose-response relationship. The data summarized in the first three chapters of the ISA allow estimates of (a) the distribution of outdoor air levels for various averaging times, and (b) the distribution of susceptibilities to specific respiratory responses within susceptible subgroups (e.g., exercising asthmatics). Based on these two components, when refined, an integrative analysis is possible that would inform the agency about the likely overall incidence of effects for current exposures, and the potential benefits of reducing exposures further. This kind of preliminary analysis would be of benefit for inclusion in a revised Chapter 4. Finally, the next version of Chapter 4 should state more explicitly whether effects are seen (or not) at current annual and 24 hr time frames and potentially at 5 minute time frames. The SO_2 level at which adverse health effects occur is a key question (stated in Chapter 1 of the ISA); it would be appropriate to directly answer this question in this Chapter. Public health impact should be addressed from a regulatory point of view and the health evidence should be discussed from this perspective.

8. What are the Panel's views on the adequacy of this first external review draft ISA to provide support for future risk, exposure and policy assessments?

In summary, this document represents a commendable first draft. A great deal of useful and policy-relevant information is presented in the carefully crafted summary statements of key findings and conclusions in Chapter 5. These kinds of summary statements provide a useful foundation for future risk, exposure and policy assessments. However, there are some areas that the committee noted that could be improved to make these summary statements even more useful for policy decisions.

With regard to exposure assessment, it is important to include cumulative exposure and spatial exposure plots of the peak 5-minute concentrations (short-term) recently collected by the states to better inform the extent to which such peaks presently occur in the U.S. It is also important to better clarify the relationships among personal, population average, and central-site SO_2 monitoring data.

In terms of collectively evaluating all considered health studies, the framework for making the evaluations of the studies needs further development, including explicitly stating the process for reviewing individual studies, the classification of the level of evidence for causality, the criteria for drawing a causal inference from the body of information in the document, and the approach to evaluating estimation uncertainty. The criteria for the inclusion of older studies (e.g., when they provide useful insights to the SOx-health relationship) need to be stated and consistently followed, as well.

With regard to deriving risk estimates from human clinical and animal toxicological studies, it is important to consider how exposure covariates might modify any SO_2-health effect relationship. For example, does the co-presence of particles or exercise alter the concentration benchmarks used in policy analysis?

With regard to epidemiological studies, it is important to better and more critically evaluate the weaknesses of multi-pollutant models.

Overall, the draft ISA covers most of the relevant evidence needed, but needs to more explicitly answer the key question "Does SO_2 cause human health effects at ambient levels?", including a more fully integrated explanation of the conclusion across the relevant lines of evidence. This will provide a more solid foundation for risk assessment and policy formulation. Specifically, the ISA needs to describe the nature of the concentration-response (C-R) relationships of the health effects with SO_2 concentration. The uncertainties of the attendant C-R response curve estimation also need to be systematically described for both the clinical studies and epidemiological studies, including a listing of their respective contributions to estimating uncertainty.

Finally, the CASAC is pleased to provide the Agency with advice and recommendations in the development of this ISA, which is a fundamental new part of the NAAQS review process. We look forward to reviewing the revised version in the coming year.

Sincerely,

/Signed/

Rogene Henderson, Chair
Clean Air Scientific Advisory Committee

Attachments

Attachment A: Roster of CASAC Sulfur Oxides Primary NAAQS Review Panel
Attachment B: Compilation of Individual Panel Member Comments on EPA's *Integrated Science Assessment for Sulfur Oxides – Health Criteria, First External Review Draft* (September 2007)

Attachment A

U.S. Environmental Protection Agency
Clean Air Scientific Advisory Committee (CASAC)
Sulfur Oxides Primary NAAQS Review Panel

CASAC MEMBERS

Dr. Rogene Henderson (Chair), Scientist Emeritus, Lovelace Respiratory Research Institute, Albuquerque, NM

Dr. Ellis B. Cowling, Emeritus Professor, , Colleges of Natural Resources and Agriculture and Life Sciences, North Carolina State University, Raleigh, NC

Dr. James Crapo, Professor of Medicine, Department of Medicine , National Jewish Medical and Research Center, Denver, CO

Dr. Douglas Crawford-Brown, Professor and Director, Department of Environmental Sciences and Engineering, Carolina Environmental Program, University of North Carolina at Chapel Hill, Chapel Hill, NC

Dr. Donna Kenski, Director of Data Analysis, Lake Michigan Air Directors Consortium, Des Plaines, IL

Dr. Armistead (Ted) Russell, Professor, Department of Civil and Environmental Engineering , Georgia Institute of Technology, Atlanta, GA

Dr. Jonathan M. Samet, Professor and Chair of the Department of Epidemiology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD

PANEL MEMBERS

Mr. Ed Avol, Professor, Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA

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

Dr. Terry Gordon, Professor, Environmental Medicine, NYU School of Medicine, Tuxedo, NY

Dr. Dale Hattis, Research Professor, Center for Technology, Environment, and Development, George Perkins Marsh Institute, Clark University, Worcester, MA

Dr. Patrick Kinney, Associate Professor, Department of Environmental Health Sciences, Mailman School of Public Health , Columbia University, New York, NY

Dr. Steven Kleeberger, Professor, Lab Chief, Laboratory of Respiratory Biology, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle

Park, NC

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

Dr. Kent Pinkerton, Professor, Regents of the University of California, Center for Health and the Environment, University of California, Davis, CA

Dr. Edward Postlethwait, Professor and Chair, Department of Environmental Health Sciences, School of Public Health, University of Alabama at Birmingham, Birmingham, AL

Dr. Richard Schlesinger, Associate Dean, Department of Biology, Dyson College, Pace University, New York, NY

Dr. Christian Seigneur, Vice President, Atmospheric & Environmental Research, Inc., San Ramon, CA

Dr. Elizabeth A. (Lianne) Sheppard, Research Professor, Biostatistics and Environmental & Occupational Health Sciences, Public Health and Community Medicine, University of Washington, Seattle, WA

Dr. Frank Speizer, Edward Kass Professor of Medicine, Channing Laboratory, Harvard Medical School, Boston, MA

Dr. George Thurston, Associate Professor, Environmental Medicine, NYU School of Medicine, New York University, Tuxedo, NY

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

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

SCIENCE ADVISORY BOARD STAFF

Dr. Holly Stallworth, Designated Federal Officer, Science Advisory Board Staff Office, Washington, D.C.

Attachment B

Comments from CASAC Sulfur Oxides Primary NAAQS Review Panel on EPA's Integrated Science Assessment for Sulfur Oxides – Health Criteria (First External Review Draft, September, 2007)

Comments from Mr. Ed Avol 2
Comments from Dr. Cowling 7
Comments from Dr. Crawford-Brown 12
Comments from Dr. Kenski 16
Comments from Dr. Kinney 19
Comments from Dr. Larson 24
Comments from Dr. Russell 26
Comments from Dr. Schlesinger 29
Comments from Dr. Seigneur 32
Comments from Dr. Speizer 35
Comments from Dr. Wyzga 39
Comments from Dr. Postlethwait 40
Comments from Dr. Gordon 41
Comments from Dr. Thurston 46
Comments from Dr. Hattis 51
Comments from Dr. Samet 59
Comments from Dr. Ultman 63
Comments from Dr. Balmes 65
Comments from Dr. Pinkerton 70
Comments from Dr. Sheppard 72

Comments from Mr. Ed Avol

General Responses to Charge Questions:
(Charge Questions 1 & 2) To what extent are the atmospheric chemistry and air quality characterizations clearly communicated, appropriately characterized, and relevant to the review of the primary SO2 NAAQS? Are the properties of ambient sulfur oxides appropriately characterized, including policy-relevant background, spatial and temporal patterns, and relationships between ambient sulfur oxides and human exposure?

The chemistry, characterizations, and properties are generally well-presented, and most of the presentation is relevant (see specific comments below for some identified exceptions). An integrated and focused summary of the current state of air quality characterization for SO2 (including what is known, what is not known, and what potential gaps need to be filled) would have been a useful addition to tie together these sections.

3. Is the information provided on atmospheric sciences and exposure sufficient for the evaluation of human health effects of sulfur oxides in the ISA?

The provided information appears to be sufficient, but additional information about intra-community variability of sulfur oxides would be helpful. It is interesting to note that although there was a previous determination that a short-term (e.g. five-minute) standard was not needed to protect the public health, there is significant effort, discussion, and monitoring based on five-minute maxima – apparently sometimes to the exclusion of monitoring for hourly or longer time-based metrics. A more productive approach might have been to monitor hourly concentrations, but retain the five-minute readings, which would allow direct comparison of multiple shorter-term metrics of interest; this should be considered for recommendation to those areas continuing to monitor ambient SOx levels.

4. To what extent are the discussion and integration of evidence on the health effects of sulfur oxides from the animal, toxicological, human clinical, and epidemiological studies, technically sound, appropriately balanced, and clearly communicated?

I was struck by the heavy reliance on (especially) clinical and (to a lesser extent) other studies from 20+ years ago, since this was supposed to be a consideration of information since the last review (ca. 1995). If little additional work has been performed but is needed, this "need" could be identified in a summary section.
Several figures compiled multiple studies into one coherent presentation, and these were especially useful. Unlike the NOx document, the SOx ISA appropriately presents informative peer-reviewed research from studies across the globe, without over-emphasizing the special status of US studies (and the ISA authors should be congratulated for this enlightened approach).

That said, Chapter Three's presentation still leaves the reader with a sense that there is a lot of information for a number of different health outcomes, but the focused integrated summary of findings, the strength of those findings, and the confidence of conclusions to be drawn from

those findings, remain elusive. Perhaps a summary table listing the various outcomes, a determination of level of confidence in the findings, and the implications of the findings would help guide the reader to a cumulative sense of the current aggregate state of SO2 knowledge.

5. To what extent does the integration of health evidence focus on the most policy-relevant studies or health findings?

The chapter in which I expected to find an integration of health evidence (Chapter 4 on Public Health Impacts) seemed to meander between a review of previous research (which would have been more appropriate for the Chapter 3 Health Effects Review) and a discussion of susceptibility (which is an appropriate topic for public health impacts, but could also have been presented in the Health Effects Review chapter and referred to from the Public Health Impacts section). I do not believe the most policy-relevant studies were well-focused into a complete integration of health evidence, although some policy-relevant studies were included. More could be done and it could be made clearer.

6. What are the views of the Panel on the conclusions drawn in the draft ISA regarding the strength, consistency, coherence, and plausibility of health effects of sulfur oxides?

The conclusions chapter (Chapter 5) begins as an extended re-visiting of previously described findings (which could be shortened or consolidated here to make the line of reasoning easier to follow). An objective, clear, short table defining the terminology of findings (e.g., causal, likely causal, inconclusive, plausible, etc) would be helpful in the document in general and in the conclusions chapter in particular. Much of the basic data needed to make a firm argument is present, but needs to be re-organized and edited. Strength, consistency, coherence, and plausibility have not been fully developed or demonstrated, but much of the data is here to do so.

7. What are the views of the Panel on the appropriateness of public health impact and the characterization of groups likely to be susceptible or vulnerable to sulfur oxides?

Comments about the discussion of public health impacts in the draft are reported in (5) above. The chapter on public health impacts contained foundational material more appropriate for earlier chapters. An important presentation about susceptibility (be it by age, genetic, or other identifying factors) was presented here, but this arguably could have been in a previous chapter and should have expanded on the inherent difference in susceptibility and vulnerability.

Background information is also included in the chapter on respiratory disease and asthma in the US, but a clear and direct linkage to SO2 public health impacts is not completely developed in the integration of this chapter (aside from the implication that if lots of disease is occurring in the US and studies have shown that SO2 causes some of it, there must be large numbers of people affected by sulfur oxides). Integration of the assembled evidence in a coherent manner still is needed.

8. What are the Panel's views on the adequacy of this first external review draft ISA to provide support for future risk, exposure, and policy assessments?

This draft is an appropriate beginning but needs some editing, revision, and re-organization to provide the needed support for future assessments. Much as the NOx ISA draft that preceded it, the SOx draft lacks a focused logical approach in which each chapter fits coherently with the next, together creating a synergistic mosaic of current scientific information, and leading to a concise and supportable number of discrete conclusions and findings. A great deal of the current draft is consumed with studies reported in the previous criteria review document or earlier, raising a question whether no new information is needed, available, or identified. There is much in the way of valuable information in the body of the document, but editing is needed to make this more fully accessible and useful.

****
****

Specific Comments:

Chapter 1

Pg1-1, lines 25-26, beginning "For the current review, multiple species of sulfur oxides..." is awkward and confusing. The following two sentences are easier to understand and explain the issue, so recommend deletion of this first sentence, and re-phasing to simply utilize the following two sentences.

Chapter 2

Pg2-3, lines 25-26 – In fact, based on recent emissions inventories (2005), Los Angeles port-related operations (primarily ship operations) account for over half of the SOx in the LA Basin.

Pg 2-3, lines 27-28 beginning "Even so, SO2 constitutes…" is interesting but irrelevant and misleading. To imply that SO2 emissions are perhaps unimportant in volcanic emissions because they constitute a minor fraction by volume ignores the fact that ground-level concentrations in the plume can be in the ppm-range of exposure. Recommend deleting this sentence and let the rest of the paragraph tell the story.

Pg2-6 to 2-7, Sources of Positive and Negative Interferences in Measurements – There is a page or more of discussion about possible interferences in the instrumentation measurement…but is this really an issue? The discussion itself points out that there are filters, scrubbers, and other commercial approaches that are routinely used to minimize these potential problems, so the reader is left with a sense of "much ado about nothing" here.

Chapter 3

Pg 3-17, lines13-32 – All of this has been covered in previous criteria documents; what is new here that justifies the authors revisiting this information?

Pgs 3-18 to 3-22 – All of this is interesting but has been previously covered in earlier reviews and criteria documents. What has been reported since the last review that would lead one to re-evaluate the current standard?

Pg 3-42 to 3-45, Section 3.1.1.7 Integration of Respiratory Effects – The discussion is useful but a bit meandering for an integrating section. Perhaps it would be useful to develop an integrated

"summary of effects" table to get a visual perception of the preponderance of evidence of effects? Such a table might list various outcomes (lung function, host defense, hospitalizations,…) and post a "+", "-", "+/-", or "?" to summarize the current state of knowledge with regard to a specific outcome. Visually, one might be able to make some judgment about the strength and breadth of available evidence for effects at a given level (which could also be presented, if desired, by listing several absolute or relative concentration columns. Absolute listings that might be considered would be actual concentrations, while relative concentration columns might be entries like "below current standard", "at current standard", or "above current standard".

Pg 3-79, line1 – Reference is made to unpublished data, which seems inconsistent with the boundary conditions of the assembled document (i.e., peer-reviewed publications since the previous criteria review).

Pg 4-2 – Most of this summation of study data seems out of place and more appropriate to the Chapter 3 presentation. This chapter is supposed to integrate the previously summarized information to focus on the public health impact, not on the reporting of individual study findings.

Pg 4-7, Section 4.2.1 Exposure of Susceptible and Vulnerable Populations – By virtue of this section title, staff has identified a potentially important perspective in the understanding of affected populations – namely, the difference between "susceptibility" and "vulnerability", in the context of ambient pollution exposures and effects. This categorization should be discussed.

Pg 4-7, lines 8-20 – This section reads more like exposure assessment than public health impact. Shouldn't this be focused on vulnerable or susceptible populations, and not digress into discussions of specific exposure sources?

Pg4-7, line 25 – Shouldn't something be added to this discussion about genetic susceptibility based on proposed mechanistic pathways of effect (such as GSTM null and oxidative stress, etc).?

Pg 4-8, lines 6-31 – Again, what is presented here is largely repetitive with Chapter 3, and more logically belongs there. This chapter should talk about public health impact, not individual study repots of observed effects.

Pg 4-18 to 4-19, lines 29 to 31 and 1 to 11 – This general health data may well be true, but it is not linked back to SO2 in this section.

Pg 5-1, line 5 in introductory paragraph – Shouldn't this be corrected to include the phrase "…since the last criteria review…"?

Pg 5-4, Section 5.2.1 Findings from the Previous Review of the NAAQS for SO2 – Why is this section here? Why not just present findings relevant to the current and recent information, which is presumably the justification for this document?

Pg 5-6 to 5-7 – These definitions of terminology (causal, suggestive, likely causal, etc) are helpful and could be placed in a table for ready reference, to clarify the increasing strength of connotation for a given word usage.

Pg 5-6, Section 5.2.2 New Findings… - This section title is inconsistent with what it goes on to contain and refer to, much of which is over 20 years old, and much of which was included in the previous cycle of document review. The focus of the current document should be the additional information available since the last document review.

Pg 5-7, Lines 29 to end – What is presented here under "New Findings of Lung Function" are old studies previously reviewed in earlier criteria review cycles. Either new studies should be reviewed, or a conclusion reached that no new studies were found, or that there is a need for new studies.

Pg5-15, Section 5.3 Conclusions, first sentence – This is what the ISA was designed to do, but in my opinion, this draft of the document is somewhat off-target.

Comments from Dr. Cowling

Very General Comments on these NAAQS Review Processes

Before dealing with the details of my specific assignment during the December 5, 2007 Peer Review of the *Integrated Science Assessment for Sulfur Oxides –Health Criteria,* I would like to offer a few general comments about these periodic NAQQS Review processes.

In a May 12, 2006 summary letter to Administrator Johnson, CASAC Chair, Dr. Rogene Henderson, provided the following statement of purpose for these periodic NAAQS review processes.

> "CASAC understands the goal of the NAAQS review process is to answer a critical scientific question: *"What evidence has been developed since the last review to indicate if the current primary and/or secondary NAAQS need to be revised or if an alternative level or form of these standards is needed to protect public health and/or public welfare?"*

During the past 18 months, CASAC has participated in reviews of three of the existing six criteria pollutants – particulate matter, ozone, and lead. CASAC has also joined with senior EPA administrators in a "top-to-bottom review" of the NAAQS review processes. These two experiences have led to a seemingly slight but important need for rephrasing and refocusing of this very important "critical scientific question:"

> ***"What scientific evidence and/or scientific insights have been developed since the last review to indicate if the current public-health based and/or the current public-welfare based NAAQS need to be revised or if alternative levels, indicators, statistical forms, or averaging times of these standards are needed to protect public health with an adequate margin of safety and to protect public welfare?"***

I hope this "critical scientific question" will be borne in mind carefully as CASAC joins with various relevant parts of the Environmental Protection Agency in completing the upcoming reviews of the primary and secondary National Ambient Air Quality Standards for Sulfur Oxides.

Thus, I recommend that every chapter in the soon to be completed Integrated Science Assessment, Risk/Exposure Assessment, and Policy Assessment/Rule Making documents for sulfur oxides (and the other five criteria pollutants) will contain a summary section composed almost entirely of a series of very carefully crafted statements of Conclusions and Scientific Findings that:

1) **Contain the distilled essence of the most important topics covered in each chapter, and**
2) **Are as directly relevant as possible to the Critically Important Scientific Question above.**

In this connection, I call attention once again to the attached *"Guideline for Formulation of Statements of Scientific Findings to be Used for Policy Purposes."* These guidelines were developed and published in 1991 by the Oversight Review Board for the National Acid

Precipitation Assessment Program. The members of the ORB who prepared these guidelines in the form of checklist questions included: Drs. Milton Russell, former Assistant Administrator for EPA, Chauncey Starr, former Director of Research for the Electric Power Research Institute (EPRI), Tom Malone, former Foreign Secretary for the National Academy of Sciences, John Tukey, Distinguished Professor of Statistics at Princeton University, and Kenneth Starr, Nobel Prize Winner in Economics. The intent of these distinguished mentors in science was to assist other scientists, engineers, and policy analysts dealing with other environmental research and assessment programs in formulating statements of scientific findings to be used in policy-decision processes. These guidelines are the best guides I know of for formulation of statements of scientific findings to be used for policy purposes:.

GUIDELINES FOR FORMULATION OF SCIENTIFIC FINDINGS TO BE USED FOR POLICY PURPOSES

The following guidelines in the form of checklist questions were developed by the NAPAP Oversight Review Board to assist scientists in formulating presentations of research results to be used in policy decision processes.

1) **IS THE STATEMENT SOUND?** Have the central issues been clearly identified? Does each statement contain the distilled essence of present scientific and technical understanding of the phenomenon or process to which it applies? Is the statement consistent with all relevant evidence – evidence developed either through NAPAP research or through analysis of research conducted outside of NAPAP? Is the statement contradicted by any important evidence developed through research inside or outside of NAPAP? Have apparent contradictions or interpretations of available evidence been considered in formulating the statement of principal findings?

2) **IS THE STATEMENT DIRECTIONAL AND, WHERE APPROPRIATE, QUANTITATIVE?** Does the statement correctly quantify both the direction and magnitude of trends and relationships in the phenomenon or process to which the statement is relevant? When possible, is a range of uncertainty given for each quantitative result? Have various sources of uncertainty been identified and quantified, for example, does the statement include or acknowledge errors in actual measurements, standard errors of estimate, possible biases in the availability of data, extrapolation of results beyond the mathematical, geographical, or temporal relevancy of available information, etc. In short, are there numbers in the statement? Are the numbers correct? Are the numbers relevant to the general meaning of the statement?

3) **IS THE DEGREE OF CERTAINTY OR UNCERTAINTY OF THE STATEMENT INDICATED CLEARLY?** Have appropriate statistical tests been applied to the data used in drawing the conclusion set forth in the statement? If the statement is based on a mathematical or novel conceptual model, has the model or concept been validated? Does the statement describe the model or concept on which it is based and the degree of validity of that model or concept?

4) **IS THE STATEMENT CORRECT WITHOUT QUALIFICATION?** Are there limitations of time, space, or other special circumstances in which the statement is true? If the statement is true only in some circumstances, are these limitations described adequately and briefly?

5) **IS THE STATEMENT CLEAR AND UNAMBIGUOUS?** Are the words and phrases used in the statement understandable by the decision makers of our society? Is the statement free of specialized jargon? Will too many people misunderstand its meaning?

6) **IS THE STATEMENT AS CONCISE AS IT CAN BE MADE WITHOUT RISK OF MISUNDERSTANDING?** Are there any excess words, phrases, or ideas in the statement which are not necessary to communicate the meaning of the statement? Are there so many caveats in the statement that the statement itself is trivial, confusing, or ambiguous?

7) **IS THE STATEMENT FREE OF SCIENTIFIC OR OTHER BIASES OR IMPLICATIONS OF SOCIETAL VALUE JUDGMENTS?** Is the statement free of influence by specific schools of scientific thought? Is the statement also free of words, phrases, or concepts that have political, economic, ideological, religious, moral, or other personal-, agency-, or organization-specific values, overtones, or implications? Does the choice of how the statement is expressed rather than its specific words suggest underlying biases or value judgments? Is the tone impartial and free of special pleading? If societal value judgments have been discussed, have these judgments been identified as such and described both clearly and objectively?

8) **HAVE SOCIETAL IMPLICATIONS BEEN DESCRIBED OBJECTIVELY?** Consideration of alternative courses of action and their consequences inherently involves judgments of their feasibility and the importance of effects. For this reason, it is important to ask if a reasonable range of alternative policies or courses of action have been evaluated? Have societal implications of alternative courses of action been stated in the following general form?:

"If this [particular option] were adopted then that [particular outcome] would be expected."

9) **HAVE THE PROFESSIONAL BIASES OF AUTHORS AND REVIEWERS BEEN DESCRIBED OPENLY?** Acknowledgment of potential sources of bias is important so that readers can judge for themselves the credibility of reports and assessments.

My Specific Assignment in this CASAC Peer Review of the First External Review Draft of the Integrated Science Assessment for Sulfur Oxides – Health Criteria

My specific assignment in preparation for the December 5, 2007 CASAC Peer Review of the "ISA for Sulfur Oxides – Health Criteria" as outlined in CASAC Chairman Rogene Henderson's memo of November 2007 is Charge Question 8 – **What are the Panel's views on the adequacy of this first external review draft ISA to provide support for future risk, exposure and policy assessments.**

Chairman Henderson also gave this same assignment to two other CASAC panel colleagues – Drs. George Thurston and Jon Samet. Thus, I am very much looking forward to comparing notes with both George and Jon during our CASAC Peer Review meeting on December 5.

My own view is that Chapter 5 of this First External Review Draft ISA – **Key Findings and Conclusions** -- is very adequate indeed in providing support for future risk, exposure, and policy assessments regarding the health effects of sulfur oxides. I offer high praise for this summary chapter because it fulfills more adequately than any Criteria Document or Integrated Science Assessment document I have seen before in providing very carefully crafted summary statements of scientific findings that conform very well to all but the last of the nine checklist questions in the above listed "Guidelines for Formulation of Statements of Scientific Findings to be used for Policy Purposes."

Each major section of Chapter 5 consists almost entirely of simple declarative statements of policy-relevant scientific findings that very adequately summarize the current scientific information contained in the earlier chapters of this ISA document including:

- Four summary statements about Emissions Sources, Atmospheric Science, and Ambient Monitoring Methods as discussed in Chapter 1,
- Five summary statements about Ambient Concentrations of sulfur oxides as discussed in Chapter 2,
- Five summary statements about Exposure Assessment as discussed in Chapter 3,
- Twenty-seven summary statements about New Findings on the Health Effects of exposure to SO_2 -- including separate summary statements derived from the scientific data and information discussed in Chapters 3 and 4 with regard to:
    - Peak (5-15 minute) Exposure to SO_2 and Respiratory Health Effects including Respiratory Symptoms and Lung Function,
    - Short-Term (24-hr average) Exposure to SO_2 and Respiratory Health Effects including Respiratory Symptoms, Lung Function, Airway Hyperresponsiveness, Inflamation, and Respiratory Emergency Department Visits and Hospitalizations,
    - Short-Term Exposure to SO_2 and Cardiovascular Health Effects,
    - Short-Term Exposure to SO_2 and Other Systemic Effects,
    - Effects of Short-Term Exposure to SO_2 and Mortality,
    - Effects of Long-Term Exposure to SO_2 and Mortaility,
    - Concentration-Response Function and Potential Thresholds, and
    - Susceptible and Vulnerable Populations.

One could of course also quarrel a bit about:

1) Whether it is optimal and correct to use SO_2 instead of sulfur oxides in very single one of the many side headings listed above.

2) Whether adequate attention is given to the chemically "reduced sulfur gases in the atmosphere" as discussed at several places on pages on pages 2-1, 2-4, and 2-24.

3) Whether it would be useful to ask each of the authors, contributors, reviewers, and EPA scientific staff to acknowledge their "professional biases" (as suggested in the last Checklist question in the "Guidelines") as well as to provide their institutional affiliations as already done on pages xiii – xx in this document.

4) Whether Chapter 2 has an optimal title. This chapter highlights key concepts or issues relevant to understanding the atmospheric chemistry, sources, exposure and dosimetry of sulfur oxides, following a "source to dose" paradigm." The idea of dealing with atmospheric chemistry all the way from emissions sources to dosimetry in the lung is a good one; but titling the chapter "Source to Tissue Dose" is a little too "cute" to be taken seriously. In my opinion, "Chemistry and Dosimetry of Nitrogen Oxides" would be better as a title for this important chapter.

5) Design and Content of Figure and Table Captions. In my opinion, every figure and table in any Integrated Science Assessment document --that is clearly to be used for policy purposes -- should "stand alone" to the maximum extent possible and not be any more dependent on descriptions in the text than absolutely necessary for understanding by readers.

6) Etc, Etc

Comments from Dr. Crawford-Brown

This review follows the Charge Questions for the document review (at the end) and provides additional comments on Chapter 5: Findings and Conclusions. A general comment is that I found the document generally appropriate, both with respect to the particular studies examined and the conclusions drawn from those studies. There was some confusion in my mind, however, about the period of time being covered. The document mentions documents produced in 1982, 1986, 1988 and 1994. It then focuses on information obtained since the 1982 document. I presume this is because it is the last year an AQCD was produced, but it seemed a long period of time to be counted as "new" data. Still, once I accepted this premise, the document flowed smoothly.

As with the NOx ISA, I am not convinced that Chapter 5 serves the purpose for which it is intended. My understanding of such a chapter is that it will provide the input into subsequent rounds of the NAAQS process. As such, it should be organized around a series of questions an assessor is likely to ask, and should provide answers to those questions so the assessor need only go back into the primary chapters to clarify some points. The answers to these questions should allow the assessor to determine whether anything about the data obtained since 1982 would cause a change (relative to the existing NAAQS) in any policy relevant question an assessor might ask in subsequent NAAQS stages. The current document does not do this. I consider the different subsections of the chapter below.

On Source to Dose Relationship, I believe the authors should have a succinct statement about whether the existing monitoring methods, including locations and numbers of samples, provide an adequate basis for estimating ambient concentrations during the period between 1982 and today. This should include a statement as to the representativeness of the monitor results for estimating ambient conditions in specific populations in the U.S., so the assessor can eventually determine which populations might be selected for analysis of scenarios of exposure. It also should include a statement of the best judgment of the ratio of ambient over personal exposures. A lot of data are presented that are relevant to this question, and they seem to suggest a ratio of about 5 or 6 with a GSD of perhaps 2, but this summary is not provided in Chapter 5. With respect to Dosimetry, the authors should provide a conclusion as to the implications of these dosimetric factors for extrapolation to the general population. If nothing else, the conclusion should be that the dosimetric results show rather clearly why the switch from nasal to oral breathing during exercise is important, which in turn alerts the assessor to the fact that the sensitive subpopulation will be people who are exercising.

On Heath Effects Findings, there should be a succinct summary of (1) the kinds of effects for which there is a relationship with current ambient levels of SOx (the authors in part accomplish this), (2) any health benchmarks suggested by the data (this is accomplished more in the body of the document than in Chapter 5, where it is most important), and (3) the concentration-response relationship obtained from the different epidemiological studies (again, the body of the document contains some excellent summary figures showing odds ratios, etc, but this level of information is not carried back into Chapter 5). As written, Chapter 5 leaves the reader with a long summary of findings without trying to focus attention onto any specific set of conclusions that can be used

by the assessor in calculating health endpoints. I realize this might be deliberate, so the ISA won't constrain the ways in which an assessor will eventually calculate risk, but there should still be some more summary statements made about health benchmarks and concentration-response functions.

Throughout this same Heath Effects section of Chapter 5, the authors use the phrases "causal", "likely causal", etc. I am supportive of this classification system, and the conclusions they have drawn seem to me justified. However, it is not clear whether these phrases are to mean something like "causal at all levels of exposure", "likely causal at all levels of exposure", etc. My concern here is that the causal link depends on level of exposure, and this classification scheme loses that distinction. It appears to be more like the practice in Hazard Identification – a practice I don't support – which asks, for example, whether a compound is or is not a carcinogen, rather than whether it is a carcinogen at specific levels of exposure and by particular routes of exposure. I think a more nuanced statement of causality is required in Chapter 5, one that focuses on whether there is a causal link at levels of exposure of policy relevance likely to be considered by subsequent analysts in the NAAQS process.

There are several places where the authors summarize a risk coefficient with a relative risk (as on Page 5-11) but don't mention that it is relative risk. If one took the first bullet in Section 5.2.2.5 literally, it would appear that a person exposed at 10 ppb would have a percent or two probability of dying! The real answer is, of course, an increased probability of dying that is one or two percent of the background probability. This is an example of where the authors of the ISA must be very careful to ensure that assessors using the document later apply the correct model of risk.

In Section 5.2.2.7, the authors go to pains to mention Krewski et al's conclusion that "the absence of a plausible toxicological mechanism by which SO2 could lead to increased mortality" suggests that "SO2 might be acting as a marker for other mortality-associated pollutants". I disagree with this statement. The lack of a mechanism being found may be simply a limitation of the existing studies. It isn't evidence one way or the other for SO2 being a marker.

Throughout the Chapter, there is no mention of the kind of concentration-response model one might expect to apply. I don't mean the shape of the model (linear, quadratic, etc) but rather the mechanistic basis. For example, the data in earlier chapters seem to suggest there is some sort of distributed threshold model at play, with each individual having a threshold but with this threshold differing from person-to-person. In that case, the shape of the curve depends on the PDF of thresholds in the population, with the low exposure portion of a population exposure-response curve being driven by individuals with a low threshold.

On Page 5-16, the authors repeat what I believe is a mistaken logic from an earlier chapter: that the public health impacts are expected to be large because the size of the susceptible population is large. A large susceptible population is important, but so is the level of actual exposures in relationship to any thresholds. I don't believe enumerating the size of the susceptible population tells us much about public health impact. Now, <u>potential</u> public health impact is another story.

Before turning to the Charge Questions, I provide here a few comments from earlier chapters.

1. In Chapter 2, some of the figures (such as 2.4-5) are almost impossible to read because the data are so overlapping.

2. On Page 2-31, line 11, I believe the authors mean Consolidated Human Activity Database.

3. On that same page, line 3, the loss processes are not only during infiltration, but due to plating on surfaces inside the structure.

4. While reading the section on Relationship of Personal Exposure to Ambient Concentration, I kept being struck by the obvious implications of these results for Tier I in the Draft Assessment Plan. These results indicate that Tier I could serve reliably as a very conservative upper bound estimate on effects.

5. In Section 2.5.4, the authors should construct a succinct statement, carried into Chapter 5, as to the implications of exposure error on slope factors and health benchmarks. They offer some good hints in this section (that random error results in bias of a slope factor towards the null, and that existence of a Personal-Ambient Ratio shifts the response curve uniformly – as mentioned on Page 2-41), but these implications are never used to draw any summary conclusions about the slope factors and health benchmarks that will appear in Chapter 5.

6. In Chapter 3, the authors provide many useful summary figures (the first is 3.1-1) but then don't provide any summary conclusions from these figures. I don't see why a summary range of odds ratios in such figures can't be established.

7. In Chapter 3, there also is clear evidence that very short-term exposures, on the order of a few minutes in exercising individuals, is sufficient to produce adverse effects at some levels of exposure. This has obvious implications for the averaging period selected in a standard, and so it should be emphasized here and in Chapter 5.

8. In Chapter 3, the authors provide some past meta-analysis results, but never attempt a meta-analysis themselves for other results, choosing instead to provide plots of the range of results. I was not sure why this is the case, and assume it is because of some mistrust of meta-analyses?

9. On Page 3-78, the authors summarize the results of the Hong Kong intervention study. They make much of the initial decline in effects after reductions in SOx, but skirt over the rebound. If the rebound shoots past the baseline mortality (and I am not saying it does as I did not find a copy of this study), this could offset the initial decline. At least some assurance that this did not take place is needed. Otherwise, the overall beneficial effect will be overstated.

10. At least on first glance, the results in Figure 3.4-1 don't appear much less conclusive than for many of the figures showing morbidity effects. I'm not sure what to draw from this observation, other than that it somewhat contradicts the position of having much less conclusive causality claims for mortality than for morbidity based solely on the epidemiological evidence.

11. In Chapter 4, when describing the epidemiological studies, I kept looking for exposure-response data plotted. It is unsatisfactory to be simply told that the data are linear without seeing them plotted with error bars.

12. At the end of Chapter 4, the authors provide an estimate of the number of individuals in the susceptible population. As I mentioned with respect to Chapter 5, the size of the population may be large but this must be coupled with exposures of health concern to produce a population risk. The latter consideration has not been applied here, although perhaps the goal was simply to show that there is a large population that is potentially at risk if exposures are sufficiently high.

I turn now to the 8 Charge Questions, drawing on many of the comments made above.

1. Yes, I feel that all of this was simple to understand and relevant. However, this is not my area of expertise, so I can't testify as to whether there might be important missing pieces of information.

2. Yes, I believe the properties are characterized at a level needed for subsequent assessments. I do, however, believe the authors need to draw a summary conclusion as to the distribution of personal-to-ambient ratios that might be applied in a variability analysis of exposure and risk. This should at least be provided in Chapter 5.

3. The information is good, but the reader is left to draw summary conclusions without adequate guidance. The authors need to provide succinct summaries of the conclusions and place these into Chapter 5.

4. The health effects discussion is clear, although again there is a need to draw better summary conclusions and point the reader towards these in Chapter 5. This is particularly true of any suggestions for health benchmarks and/or exposure-response relationships.

5. The document does focus on the most policy-relevant findings, but (as mentioned in Question 4) does not draw summary conclusions that will be needed in policy determinations. This is in part because the authors have chosen a system of causal claims (causal, likely causal, etc) that does not specify the exposure level at which these claims are valid. This aspect needs to be improved.

6. My answer here is the same as in Question 5. The causal claims are justified, but need a bit more nuance by stating the levels of exposure at which they apply.

7. The appropriate subpopulations have been selected, but the document gives the incorrect impression that a large number of individuals in this subpopulation means a large public health impact. This is not true unless these subpopulations are also exposed at levels above health benchmarks.

8. With the improvements I have suggested for Chapter 5, the document can provide an adequate basis.

Comments from Dr. Kenski

General: Kudos to the team preparing these documents! This ISA was in much better shape than the NOx ISA we reviewed a month ago. I found it to be well written and thoughtful. It was also more truly integrated, in the sense that there were useful summaries that drew reasonable and defensible conclusions, without overinterpreting or putting too much spin on the data.

Section 2: The description of SOx chemistry is adequate. However, the discussion of ambient air quality data for SO2 would benefit from some additional detail. For example, the Figures 2.4-2 and 2.4-3 give a good regional picture, showing the decline in regional SO2 and SO4 over the last 10-15 years. But because this is CASTNET data, it comes from monitors that are sited in rural areas that are, by definition, far from sources and people. So these plots probably don't show us concentrations that the majority of people are exposed to. Wouldn't a map that used the AQS data, including all those urban sites, also be informative? A map of the NAMS/SLAMS network should be included here or in the Annex. We are given Table 2.4-2, but it never actually says how many monitors are providing data – is it the sum of the 3-yr averages inside and outside CMSAs? Are these monitors just NAMS/SLAMS, or are CASTNET sites included? Much is made of the fact that the CMSAs with 4 or more monitors have plenty of data below the detection limit, but I don't believe it's ever mentioned that concentrations are higher in urban areas than in rural areas, at least as evidenced by the medians and 75th %iles in Table 2.4-2. Surely this is important to our basic understanding of SO2 ambient concentrations.

p. 2-13: Figure 2.4-5 is not very helpful as currently plotted; perhaps using a log scale would be more effective at showing what the real distributions are. Lines 11-13 on p. 2-13 state that (using this plot as a basis, presumably) highest concentrations are reached at midday or during the middle of the night. That may be true, but this figure does not really allow one to draw that conclusion. If this is a point that should be made, then a better graphic is required, one that draws our attention to the central tendency of the data and not those pesky outliers.

p. 2-14: The lack of correlation among SO2 monitors was helpful to point out; it would be also be helpful to augment this discussion with a summary of how the monitors are sited (as described in the AQS system at a minimum)—i.e., how many are source oriented, how many are community monitors, even if it's relegated to the Annex. Because some monitors are source oriented and others population-based, it is perhaps not surprising that there is a lot of intra-city variability. A review of data that show this intra-city variability (e.g., saturation studies) would be useful and would help readers understand that monitors in the national networks have a limited ability to characterize the concentration gradients that exist in urban areas (for a very recent example, see Wheeler et al., Intra-urban variability of air pollution in Windsor, Ontario—Measurement and modeling for human exposure assessment, Environmental Research 106 (2008) 7–16, available online).

p. 2-24: The discussion of correlations between SO2 and SO4 needs a conclusion, even if it's as simple as noting that there is no consistent evidence for correlations across the country. This section also needs some quantitative description of correlation or lack or correlation between

SO2 and other copollutants, especially in light of the health effects that can be confounded by these.

Section 3: I appreciate the consistent use of summaries at the end of each of these sections on health effects.

p. 3-2, lines 1-15: In light of the data presented in Sec 2 that showed no or very weak correlations between SO2 and SO4, this paragraph is confusing. Multipollutant models should perform best when the species of interest are not correlated. Effects of SO2 might be confounded by copollutants, but SO4 seems unlikely to be one of them.

p. 3-45, line 8: should this be 'There were also *no* key human…?

p. 4-1, line 31: needs a comma after effects
line 32: needs a comma after background level
p. 4-10: lines 1-2 are repeated from previous page, and lines have apparently been dropped from the bottom of this page

p. 5-2, last bullet: This bullet should make it clear that it's summarizing concentrations that are averaged over these regions. Because the NAAQS applies to individual monitors, not to regions, another bullet should summarize the actual NAAQS-relevant concentrations, i.e, give the ranges of annual average and max 24-average concentrations at the regulatory monitors (e.g., the interquartile range of annual average concentrations was from 1 to 6 ppb in urban areas, and the maximum annual average was 148 ppb at Some city, Some State.) The conclusions put it perfectly (p. 5-15, lines 22-26) although, oddly, these same numbers aren't presented anywhere else in this Section.

p. 5-3, line 1: This bullet is a bit of a red herring. The generally slow conversion of SO2 to SO4 and the fact that emissions are often from hot plumes and elevated stacks are sufficient reasons to believe that the 2 species would not be highly correlated at ground based monitors. The more important conclusion from the correlations presented in Sec. 2 is that there is huge spatial heterogeneity in SO2 on an urban scale. This finding is the one that has the most serious implications for exposure assessment, because it means it will be more difficult to accurately characterize the concentrations that populations are exposed to.

As mentioned above, we should have some a bullet here about what other pollutants SO2 is correlated with, if any.

p. 5-9, line 32: amount -> among

Annexes: These really need a comprehensive table of contents that includes subheadings so those unfortunate souls reading it without an electronic version can actually find information without thumbing through a hundred pages or so.

Annex 2: It does make sense to include NOx chemistry in this ISA, but it was a little disconcerting to open up Annex 2 and find the first 20 pages devoted to NOx instead of SOx.

Please add a prominent sentence or two to the introduction explaining the rationale for this decision, as it is easy to overlook the footnote on p. 1-5. The aforementioned table of contents would make the logic and presentation of this information more apparent as well.

Table AX2.6-1 is not very clearly formatted. It's difficult to determine which items are summed to make up the subtotals. Offset the totals from the individual entries or otherwise make it obvious that not all of the numbers in each column can be added together. Some items look like they should have been bolded but are not – e.g., solvent utilization? Metals processing?

Comments from Dr. Kinney

Specific Comments:

P. 2.10, figure 2.4-2: add footnote for conversion of ug/m3 to ppb concentration units

p. 2-31, line 18: change "characterized" to "simulated." It is important for readers to understand that you are not estimating actual personal exposures of individuals by these methods, but you ARE simulating the population distributional characteristics.

p. 2-32, para ending line 31: Add the limitation that, because of sensitivity and LOD issues, it is impossible to characterize hourly or shorter personal exposures, which is a major limitation of available personal monitoring technology since micro-environmental exposures at these time scales are likely to be very important.

p. 2-37, para ending line 3: I question whether these studies warrant such extensive and detailed review here given the LOD problems with the personal samples. I would replace this with a short paragraph stating that LOD issues with currently-available technology preclude our ability to address ambient/personal relationships for short averaging periods. The paragraph which follows is probably adequate. Just refer further discussion to the Annex.

p. 2-38, line 14: This is an odd way of expressing general concerns over confounding by co-pollutants, and it coveys a strong presumption that SO2 cannot be the true causal factor. Please edit to offer a more balanced viewpoint.

p. 2-42, line 9: I don't follow the logic here.

p. 2-45, whole section: Throughout this section, this discussion of dosimetry provides far too much detail about studies, most of which were presumably reviewed extensively in previous CDs. Instead, here it would be sufficient to summarize the common findings regarding regional dosimetry with and without exercise. Details are not needed.

p. 2-46, line 16: this summary paragraph should be the main content of this section, as noted above.

p. 3-2 line 11: this statement about SO2 and SO4 is directly contradicted by the data presented in chapter 2 demonstrating very low correlations between SO2 and SO4 in most locations examined.

p. 3-2, line 13: It would be fair to say that two-pollutant models involving both SO2 and PM2.5(or so4) are a valid way to assess the relative health impacts of the two pollutants. If one or the other pollutant is more robustly associated with health, that supports the interpretation that the robust pollutant may be the more valid measure of risk.

p. 3-6, first paragraph: Were any of the co-pollutants independently assocaited with symptoms, and if so, were their effect estimates as robust as was SO2's? See lines 29-30 of same page for the kind of information that is needed here.

p. 3-9, lines 17-18: This possibility seems to be raised out of context. The question is whether positive studies for SO2 have examined robustness wrt PM. If they haven't done so, or SO2 is generally not robust when they did, then it's reasonable to raise this concern, but absent a link to the findings of the studies, this is a red herring that conveys an a priori bias towards PM effects on the part of the author. In the text that follows, the only study that actually examined this issue found little evidence to support the statement.

p. 3-16, line 2: O3, not PM, was generally the most robust pollutant in the health studies reviewed above.

p. 3-22, line 28: this is very light exercise

p. 3-23, second para: The Koenig results at 0.1 ppm should be noted here as well, since the co-exposure regimen used there is a realistic simulation of ambient conditions.

p. 3-27, first full para: Did Boezen examine co-pollutant effects in either study? Worth a mention here.

p. 3-27, second full para: what concentration of SO2 was used in the Nowak study?

p. 3-30, line 25: this is confusing. Pneumonia isn't part of COPD. Edit.

p. 3-31, lines 9-11: This statement is not supported by the evidence you've presented, e.g., Wilson et al results for ages 15-64. This is important as this concept of no adult effects is carried through the rest of the document. In figure 3.1-8, results jump around a lot for all ages, no more so for the adults than any other age group in my view.

p. 3-35, line 13: Seems inconsistent with Petroeschevsky results plotted in fig 3.1-7

p. 3-35, line 18-20: Needs re-phrasing. Say something like many studies observed positive associations, some of which were statistically significant.

p. 3-35, lines 25-29: among the few adult results, it's true that few are statistically significant, but Wilson is, and several others are consistently positive. There's really no basis to make claims about effect modification by age based on the available data displayed in figure 3.1-10

p. 3-39, lines 1-3: this is a more accurate summary than the one that led off the section.

p. 3-40, lines 15-17: My read of figure 3.1-11 is that SO2 has sometimes been robust and other times not in co-pollutant models. This summary statement, and the similar one leading off the paragraph, is not consistent with the evidence.

p. 3-42, line 5: replace “generally robust” with “often robust”

p. 3-44: first para: Should mention the Nowak et al. human controlled exposure study in this paragraph too. It seemed like the most clear indication that BHR was associated with SO2 response.

p. 3-44, line 27: change “were not sensitive” to “were moderately sensitive”. Also, delete second half of sentence.

p. 3-45, lines 1-2 (and previous lines): I find this statement about biological plausibility to be unwarranted given the hugely different SO2 concentrations used in the epi vs. experimental studies.

p. 3-45, line 8: Did you mean to say “there were also no key..”

p. 3-45, line 15: add a sentence explaining what these markers have to do with cardiovascular health.

p. 3-46: The reader does not need so much detail on these studies of HRV since they say very little about SO2 effects per-se. Reduce to one short paragraph summarizing overall findings and interpretational problems wrt SO2.

The same is true of section 3.1.2.3 on cardiac arrhythmias. The epi studies reviewed there tell us little or nothing about SO2 because of the co-pollutant mix. I would eliminate most of the detail and simply state in a few sentences that there is some evidence for SO2 but that it's confounded by co-pollutants that are likely more relevant, i.e., PM.

p. 3-50, section 3.1.2.4: these first two paragraphs could be reduced to one sentence stating that there is no clear epi evidence for robust SO2 effects on blood pressure.

p. 3-51, lines 12-13: this says it all, and is all that needs to be said. The annexes are where detailed study reviews need to be.

p. 3-51, section 3.1.2.5: same thing

p. 3-53, para starting line 14: what happened when other pollutants were added in these studies?

p. 3-57, section on cerebrovascular effects: I wonder whether it is necessary for this document to catalogue every published health outcome for which SO2 results have been reported, regardless of how biologically implausible or far-fetched the cause-effect relationship is. Just because it's been reported doesn't mean it meets the standard of relevance to understanding the health effects of SO2. The limitations of epidemiology, especially ecologic time series or cross sectional studies, adds another major layer of uncertainty, rendering these findings largely irrelevant to the current purpose.

p. 3-58, lines 7-8: While I tend to agree that any observed SO2 effects seem likely to represent some form of confounding, the results in figure 3.1-13 do not support this statement.

p. 3-60: Exposure levels and/or durations are so far from ambient in most of these experimental studies that I question the inclusion in this document which is dealing with setting an ambient air standard.

p. 3-62, lines 16-20: this is all that needs to be said about this literature. just add the relevant citations to this short paragraph and delete above text and table.

p. 3-64, line 24: which is a good thing in my view. the BAD thing about meta analyses wrt multicity studies is the problem of publication bias.

p. 3-75, line 17, insert “to some extent” between “confounded” and “by”

p. 3-79, sentence from line 15-17: delete. Not relevant.

p. 3-80, lines 5-6: I agree with this comment about biological plausibility, but it appears denovo here, without any thoughtful discussion about biological mechanisms and plausibilty. A section should be added which provides such discussion if this statement remains here. Otherwise, can do later in an integrated assessment.

p. 3-82, line 8: what about patterns of co-pollutant concentrations and health effects?

p. 3-82, line 18: delete “moderately”

p. 3-84, line 13: Add that a major problem with geographic studies comparing several different communities is confounding by co-pollutants. There are few if any situations where SO2 is the only pollutant that varies across metro areas. At best, such studies may indicate something regarding health effects of "bad air" but rarely will provide pollutant-specific information. In light of this, I think the detailed study descriptions given here are more than is necessary and the whole section could be reduced to one long paragraph with references.

p. 3-86, line 8: In addition, level and duration of exposure were far larger than ambient conditions.

p. 3-88, line 13-14: once again, need to mention the very much higher than ambient concentrations used here.

p. 3-93, line 16 and others: this is not air concentration units; what does it mean?

p. 4-1, line 28: shallow slope not really an “error source”

p. 4-3, first para: aren’t the Koenig et al., ozone/so2 results relevant here (0.1 ppm SO2 effects)?

p. 4-14, lines 7-8: as noted earlier, i'm not sure this exclusion of intermediate ages is justified by the evidence.

4-14, lines 26-27: what is meant by "reduce expression of function in the lung?"

p. 4-16, line 23; delete “other”

p. 4-18, line 11 e.g.,: there should be a greater focus here on asthma as the primary disease condition of interest with respect to SO2.

p. 4-19, last para: It would be helpful to summarize population numbers for these specific groups here, both as numbers and proportion of total population.

p. 5-1, line 31, append: “As a result of this chemical transformation, SO2 concentrations diminish downwind of sources as sulfate concentrations increase.”

p. 5-2, line 20, insert “of hourly concentrations” after precise measurements..

p. 5-8, line 20: change “generally” to “often”

p. 5-10, line 1: change “generally” to “often”

p. 5-10, lines 2-6; I don't think it's reasonable to invoke biological plausibility when concentrations are two orders of magnitude higher in the experimental vs. epi studies. Need more nuanced statement here that takes this uncertainty into account.

p. 5-10, line 27: change “of” to “on”

p. 5-14, line 18: change “respiratory diseases” to “asthma”

p. 5-14, lines 23-24: change “respiratory illnesses, particularly asthma” to “asthma”

p 5-15, line 3: here’s another statement about ages, which is only weakly supported by the evidence in my opinion.

p. 5-16, line 5: insert “somewhat” before robust.

p. 5-16, line 8: delete phrase after “causes”

p. 5-16, line 10: delete “these”

p. 5-16, line 19, insert “,difficulty in separating SO2 effects from other co-pollutants,” after “risk estimates”

p. 5-17, line 16-21: It seems that the Koenig et al., o3/so2 controlled exposure study would be relevant to mention here, or at least the observation that ozone appeared to potentiate so2 effects.

Comments from Dr. Larson

General Comments on Chapter 2 Charge Questions

1. To what extent are the atmospheric chemistry and air quality characterizations clearly communicated, appropriately characterized, and relevant to the review of the primary SO_2 NAAQS?

This chapter has an adequate but succinct summary of the relevant atmospheric chemistry of SO_2 as it relates to the ultimate levels of SO_2 in the ambient air. The air quality characterizations are also clearly communicated. However, due to the preponderance of below detection values in the regulatory observations, it is difficult to get an accurate picture of the actual concentrations in most urban areas. For example, to assess the spatial variability across metropolitan areas there are only 12 MSAs with four or more monitors, with some of these located near industrial sources. Contrast this with the SAVIAH epidemiological study by Pikhart et al,(2001) where outdoor passive samplers were deployed at over 100 sites to estimate spatial variation.

While it is true that more and better measurement methods would provide more precise characterizations (for a more sensitive method, see Matsumi et.al, Atmospheric Environment 39 (2005) 3177-3185), it is not clear that the focus should be on a broad regional characterization. It would seem just as reasonable to focus on those areas with relatively high SO_2, either selected MSAs or locations near major sources. One might then provide more a more thorough air quality characterization in a more limited geographical region. From an epidemiological perspective, the relevant co-pollutants may be different in these higher SO_2 regions than in the country at large or in the European and Asian cities where adverse health associations are observed.

2. Are the properties of ambient sulfur oxides appropriately characterized, including policy-relevant background, spatial and temporal patterns, and relationships between ambient sulfur oxides and human exposure?

The characterization of policy relevant background levels is mostly appropriate. However it is a little misleading to say that these levels can comprise up to 70% of the SO2 in the northwestern U.S. Most urban areas in this part of the country are not strongly influence by volcanic emissions.

I would also take exception to the conclusion on page 2-37 that "when personal exposure concentrations are above detection limits, a reasonably strong association is observed between personal exposures and ambient concentrations". This statement is essentially repeated in the conclusions chapter in section 5.1.3. As best I can tell, this is based primarily on the results of Brauer et al., 1989. While this might be a perfectly good study, it was done in one city over two seasons. One study does not justify this important generalization.

Is the information provided on atmospheric sciences and exposure sufficient for the evaluation of human health effects of sulfur oxides in the ISA?

From an epidemiological perspective, there is not sufficient information. The current ambient levels are very low and poorly measured in most locations. There is also insufficient information to predict personal exposure levels from these low ambient levels measured at most community monitors. From a toxicological perspective, perhaps one can draw conclusions in a limited number of locations based on the short-term effects seen in controlled human exposure studies and the measured hourly or daily ambient levels at the upper end of the concentration distribution.

Specific Comments on Chapter 2

page 2-13 The mean values are visually difficult to distinguish from those above the 95^{th} percentile in Figure 2.4-5.

page 2-14 Are the correlation coefficients based on hourly or daily values?

page 2-30 Lawn equipment?? Perhaps this refers to PM or VOC emissions.

page 2-35 Is there a reference for the general statement that SO_2 concentrations increase with height above ground? This may be true near elevated point sources, but not during nighttime inversions.

Page 2-36 What can we conclude from these negative slopes? That personal exposures decrease with increasing ambient SO_2 concentrations??? This discussion is confusing.

Page 2-28 The last sentence starting with "Thus," is poorly worded.

Comments from Dr. Russell

Like the NO2 ISA, I like the idea of the second chapter being nicely trimmed down, getting to the point of going from the source to the dose. Similar to before, there are areas to be strengthened and refocused.

First, the discussion of the gas phase oxidation of SO2 to sulfuric acid, and then what happens to sulfuric acid monomers needs to be strengthened. Sulfuric acid monomer is, as stated, very water soluble. However, upon formation, its very low vapor pressure is rather key. It will condense on preexisting particles or can nucleate to form nanoparticles. It will not only transfer to water droplets. Further down page 2-2, (line 25), I would say that at pH's above 5.3, ozone oxidation becomes increasingly important. Another piece of chemistry to be brought in to this discussion is that as sulfuric acid is formed, the pH of the aerosol drops, and this may lead to increased formation of VOC. Perhaps this is to be covered under the PM review.

On page 2-3, line 16, one should note they are referring to US emissions. Lines 19-23 are a bit confusing at present. What is important, here, is that virtually all of the fuel-bound sulfur gets oxidized to a volatile component (SO2 or SO3), and that there is virtually no sulfur in air, so the sulfur emitted from burning a fuel is quantitatively related to that in the fuel.

I particularly think that the section on Measurement Methods needs to be refocused. The major question to be addressed here is if the current methods employed in the field provide reliable measurements of SO2 for levels of interest, and this should be answered quantitatively. At present, there is significant discussion of the various measurement approaches (a little is needed) and lots of discussion on possible interferences, but never does one get the answer to what is the typical uncertainty in the measurements at a typical monitor in the US. I suspect that the method employed, while subject to some interferences, provides perfectly fine data, and that the level of uncertainty is such that we need not concern ourselves with possible interferences and biases. Quantify the problems, let the reader assess if they are of concern.

Figure 2.4-5 "… in focus" in focus of what? Actually this figure is not overly instructive since most of the data is very much at the bottom end, and it needs to include more information (what years…).

In considering emissions, it would be good to also provide some information as to future emissions for perspective. CAIR is going to significantly lower emissions in areas where they are currently high. This is important for our further consideration as to how a standard might impact air quality.

What are the PRB levels of sulfate? What are the PRB levels of deposition? I am not sure this has to be added here, as part of the Sox primary document, but the section title suggests it will be.

Page 2-23, line 12: I think you mean months, not seasons. (or does Philadelphia have three summer seasons… which in some places might be nice, but having spent time in Philadelphia in the summer, I am not sure it is good there.)

The Findings and Conclusions Chapter is still rather rough. I would also focus on what parts of the prior chapters significantly impact how the NAAQS may be revised. SO2 is a slightly soluble gas, and the fraction that is oxidized in the aqueous phase is quite dependent upon location. I am not sure what is meant by quantitatively on page 5-1, line 30. (Both lines 23 and 30 should be reworded.) Section 5.1.2 should also look to the future, given CAIR. It is interesting that the ISA says that it is inadequate for measurements at or below 3 ppb, but then notes that the average on the West Coast is ~1 ppb.

Post meeting additions:

While suggested above, I want to re-emphasize the need to better characterize the potential for significant exposure measurement error with SO2. This is strongly suggested by the Sarnat et al., results, but the is significantly more at issue than might be appreciated based upon that work alone. Exposure to high levels of SO2 will come predominantly from being in the plume of a point source that burns coal or oil. A direct mass balance suggests that a high observation will be from a plume that is rather thin, thus impacting a rather small portion of an area. On the other hand, that plume, on a day that has a wind blowing at a slightly different direction will miss the monitor, but impact a similar area, and likely a similar population. Chapter 2 needs to address this in more detail, and Chapter 3 should discuss the epidemiologic studies in light of how well they have addressed this potential error, and also the co-exposures. Further, if one looks at the possible concentration gradients in plumes, one can see that a near-hit can have a significantly different concentration than a direct hit.

In an effort to help assess how to extrapolate the information we currently have to better understand the distribution of pollutant concentrations at shorter averaging times, it would be good to provide, as the data allows, CDFs of 24 and 1 hr, and 5-minute, concentrations, and correlations between 1-hr and max 5-minute average concentrations. I might also include a box-and-whisker plot of 5 min concentrations, possibly by hour of day.

To better assess the possibility of confounding, it would be good to provide information as to how SO2 correlates with concentrations of other pollutants, including metals.

In regards to the points brought forward to Chapter 5, I might suggest that the bullets need to be modified. First, the route of SO2 oxidation is heterogeneous spatially. I would not call gas phase oxidation of secondary importance (I would also suggest that one try to use results from a more detailed, US-specific model, e.g., CMAQ, if such an analysis exists).

I take issue with the second bullet in 5.1.2: The LOD may vary, and I suspect the various QA tests support that the monitors in many locations are good down to 1 ppb. Don't be so sweeping in the bullet. Likewise, in the fourth bullet, I would suggest that there is little (not no) correlation.

5.1.3 Needs to have a MAJOR bullet on the issues with exposure measurement error. It is potentially large, and needs more work throughout the document. The bullet should

specifically note that exposure measurement error is likely larger for SO2 than other pollutants. Further, this is potentially larger as the reporting time gets shorter.

Comments from Dr. Schlesinger

Overall, this is a very well written document.

Chapter 3 is especially clear, in that each section has at the end a concise summary indicating the overall conclusions from the studies discussed.

p. 1-1, line 24. Put ")" and then "and" after]

p. 1-4, line 11. The lab studies in the document are not all at or near ambient SOx levels.

p. 2-42, line 23. Add "and systemic" after "…respiratory tract"

p. 2-42, line 24. Add "irritant" after "epithelial"

p. 2-42, line 32. Add "respiratory functional parameters, e.g….." after layers;

p. 2-43, line. 20. "sooner" is not really a good scientific term to use here.

p. 2-43, line 24. Add "at ambient concentrations" after "…at rest." This makes it more consistent with p. 2-45, line 5 that indicates that absorption is related to concentration.

p. 3-55, line 15. Delete "…the collective evidence that…" and replace with "…any association between…"

p. 3-55, line 16. Delete "…has an effect of…" and replace with "…and…"

p. 3-62, line 20. Define what is considered to be a "high concentration."

p. 3-64, line 21. What were the criteria for selecting "some" of the studies.

p. 3-68, line 20. The differences between the constituents should be elucidated.

p. 3-74, Section 3.2.1.4. The potential confounding by copollutants has been presented in prior sections of the chapter related to health outcomes so it is not clear why this section is needed. It is somewhat redundant.

p. 3-75, Section 3.2.2. Material presented here has been discussed in earlier sections of the chapter.

p. 3-81, line 1. Add the word "…subtle changes in…" before cilia. Some changes could be seen with light microscopy.

p. 3-84, line 17. Be more specific about the aspects of respiratory health affected. For example, the sentence could read, "…health, such as chronic bronchitis, asthma or respiratory symptoms."

p. 3-94, Section 3.4.1. Some of the results in thesestudies related to carcinogenesis should have been discussed in Section 3.3.2.

p. 3-103, line 14. Delete "…sulfur agents…" and replace with "…particulate sulfur oxides…"

p. 3-104, lines 3-4. This sentence does not really indicate what the association is with in terms of health outcomes.

p. 3-104, line 9. After "…confounding…" add "…and lack of underlying biological plausibility…"

p. 4-6, line 14. Add "…measured…" after "…entire…" and then add "…over the measured concentration range." After "…effect…"

p. 4-7, Section 4.2.1. This section is not very coherent. The discussion presented should be melding within other relevant sections.

p. 4-8, Section 4.2.2.1. Details of the studies presented here should have been described in Chapter 3. Chapter 4 should ideally be an integration and synthesis chapter taking the material from prior described studies to provide a conclusion as to public health impact.

p. 4-12, line 17. Add "…or a general reduction in immune competence." After "…network…"

On page 4-12, lines 8-9, it is noted that two susceptible groups are children and infants. On page 4-14, lines 7-11, it is noted that there is limited evidence that children are more susceptible to SO2. This is somewhat of a contradiction and should be addressed. Similarly, in Section 4.3, it is noted on page 4-16, lines 15-16, that exposure to SO2 is associated with various outcomes particularly among asthmatic children. Again, there is the appearance of contradictions in the discussion presented. This issue arises again on page 4-17, lines 24-25 and line 32 and page 4-19, lines 11-18. There needs to be more consistency in the issue of susceptible populations.

p. 5-4, lines 17-18. It is stated here that the thoracic region is more sensitive than the upper airways. However, there are receptors in the upper airways that could trigger an asthmatic attack.

p. 5-6, line 24. After "effects" add, depending upon the relationship between exposure concentration and actual ambient levels."

p. 5-7. Section 5.2.2.1. In this and other sections, the term "peak" appears to refer to duration of exposure rather than actual concentration. There may be some confusion among readers since many will relate peak to concentration and not time.

p. 5-9, lines 13-14. Change sentence as follows: "…biologic plausibility, but no concentration-response information to allow a mechanistic understanding of…"

p. 5-9, line 32. "amount" should read "among"

p. 5-10, line 27. For consistency, the word “weak” should be replaced with “inconclusive.”

p. 5-16, line 30. Fog droplets are usually large and therefore would not be carriers of SO2 to the distal airways.

p. 5-17, line 15. After “… mixture…” add, “or other chemical component within it.”

p. 5-17, line 21. After “metals” add “on particulate matter”

Chapter 5 should be retitled, Summary and Conclusions.

Comments from Dr. Seigneur

Chapter 1, Introduction

Charge question 3: This chapter states that all sulfur oxides will be treated in the ISA. Yet, the emphasis in the rest of the document is clearly on SO_2. EPA may want to consider whether it may be more appropriate to mention in the Introduction that although all sulfur oxides must be addressed, the ISA focuses on SO_2 because (1) sulfur oxides in the atmosphere are mostly SO_2 and sulfate and (2) sulfate is treated under the PM NAAQS.

Chapter 2. Source to tissue dose

Charge question 1: The discussion of the atmospheric chemistry of sulfur oxides covers all the important points but some revisions would help make the discussion more precise.

Page 2-1, lines 24-25: Replace "partitions into the aqueous phase of particles" by "forms new particles by nucleation or condenses on existing particles" (note that under dry atmospheric conditions, H_2SO_4 may form non-aqueous ammonium sulfate or bisulfate).

Page 2-2, lines 3 and 4: The same logic applies here, under dry atmospheric conditions, H_2SO_4 may not be transferred to an hydrometeor (since there may be none) but instead to a dry solid particle. Also, solubility in water is not the main reason why H_2SO_4 is transferred to the particulate phase (HNO_3, which is also very soluble in water, tends to remain mostly in the gas phase because it has a high vapor pressure); the very low vapor pressure of H_2SO_4 is the driving force for its rapid transfer from the gas phase to the particulate phase, regardless of the ambient relative humidity or ammonia concentrations. I suggest replacing the existing sentence by the following one: "Because H_2SO_4 has a very low saturation vapor pressure, it will be rapidly transferred from the gas phase to a condensed phase (particulate matter or droplet).

Page 2-2, line 11: I believe that manganese (Mn) is also a catalyst for the oxidation of SO_2 by O_2 in aqueous solution. Is Cu an important catalyst compared to Fe and Mn?

Page 2-2, line 14: Add Finlayson-Pitts and Pitts as one of the standard references for atmospheric chemistry.

Page 2-2, line 24: Why a pH of 5.3 ($[H+] = 5 \times 10^{-6}$ M), and not 5 or 5.5?

Page 2-3, line 3: A sentence needs to be added to highlight the fact that the oxidation of SO_2 by O_3 and O_2 is self-limiting, because as sulfate is formed, the pH decreases and, consequently, the kinetics of those reactions decrease.

Page 2-3, line 9: I think that this example of (gas-phase) SO_2 oxidation in power plant plumes is misleading because SO_2 oxidation in power plant plumes varies from nearly zero to a value that may exceed the rate of oxidation in the background air. Near the

stack, the high NO concentration in the plume depletes the oxidant concentrations (including OH) and, therefore, there is almost no SO_2 oxidation in the gas phase. Farther downwind, the oxidation of SO_2 will pick up as the NO concentration is diluted with background air. Then, in a NO_x limited environment for oxidant formation, the oxidant concentrations in the plume will exceed those in the background air and, therefore, the SO_2 oxidation rate will exceed that of SO_2 in the background. A range of 0.5 to 2% fails to reflect the complex evolution of the SO_2 oxidation process in power plant plumes. I suggest stating that SO_2 oxidation in background air is on the order of 1% per hour; then, adding a sentence explaining the three stages of SO_2 oxidation in power plant plumes.

Charge question 1: Although it is stated in Section 1 that this document addresses all sulfur oxides, Section 2.2 on the sources of sulfur oxides does not address sources of SO_3 and sulfate. Although emissions of sulfate from industrial processes are typically a very small fraction of SO_x emissions, it would be useful to mention those and to indicate the fraction of SO_x emissions that is sulfate for major source categories (e.g., coal-fired power plants, diesel engines).

Charge question 1: Page 2-9: It would be useful to discuss briefly the issue of Sulfur Emission Control Areas (SECAs), i.e., those areas where ocean-going ships will have to burn lower-sulfur fuel to minimize their impacts on air quality inland. Ship emissions are a major issue in ports and along shipping lanes and as sulfur emissions from point sources and on-land mobile sources decrease, ship emissions take more relative importance.

Pages 2-10 and 2-11: I think that CASTNet is now written CASTNET.

Page 2-11, line 1: Spell out CONUS as continental (or contiguous) United States.

Page 2-23, line 25: Why does SO_2 peak during summertime in Los Angeles?

Charge question 2: Page 2-27, line 24: Since it is stated in Section 1 that this document addresses all sulfur oxides, some discussion of the policy relevant background (PRB) for sulfate should be added.

Chapter 5. Findings and conclusions.

Charge question 1: Page 5-1, line 30: delete "in cloud drops and/or in particles" because sulfate is also formed in the gas phase. The intention was perhaps to state that sulfate ends up in drops or in particles, then add "; sulfate is then transferred to droplets or particles due to its very low saturation vapor pressure".

Page 5-2, lines 18-24: It is good news that better detection limits will be made available for routine monitoring of SO_2. Can EPA give an approximate timeline for the availability of those new monitors in the routine network?

Page 5-15, lines 24 and 46: Did you mean <120 ppb and <600 ppb?

Annex 2.2. Chemistry of nitrogen oxides in the troposphere

Is this annex needed in the Sulfur oxides ISA? If so, the comments that I provided earlier on the Nitrogen oxides ISA apply here (in particular, revisions to Figure AX2.1-1).

Annex 2.3. Chemistry of sulfur oxides in the troposphere

Page AX2.24, lines 21-22: Although the solubility of H_2SO_4 is relevant for its removal by droplets, it is not very relevant to its gas-to-particle conversion; its low saturation vapor pressure needs to be invoked here.

Annex 2.4. Mechanisms for the aqueous formation of nitrate and sulfate

Page AX2.28, line 4: Include manganese as one of the catalysts of aqueous SO_2 oxidation.

Annex 2.7. Methods to calculate concentrations of nitrogen oxides in the atmosphere

Is this section needed in the Sulfur oxides ISA? Also, there should be a similar section on the topic of modeling sulfur oxides, in particular SO_2. A discussion of AERMOD is warranted.

Annex 2.9. Policy relevant background concentrations

Table AX2.6-1. Do those emissions include emissions from ships in coastal waters and/or Sulfur Emission Control Areas (SECAs)?

Comments from Dr. Speizer

Pre-meeting Comments on: Integrated Science Assessment for Sulfur Oxides First External Review Draft September 2007

Charge Questions (paraphrased):

1-3. Atmospheric Chemistry and air quality characterization appropriate and relevant for review of primary SO2 NAAQS, properties of ambient SO2, and relevant special, temporal patterns and exposure estimates.

Page 2-9 Figure 2.4.1. Better resolution of this figure is needed. In fact, it is not clear what state summarized values mean. Surely the emission inventories for a state like Ohio are dominated by point sources in contrast to a state like Montana or even Pennsylvania. The text does a little better job by suggesting that there are localized regions, but the Figure is misleading. Subsequent figures are better.

Page 2-23, end of first para at lines 6-7 and contrast with remaining two paras. This is a curious statement, given the apparent inverse correlation seen in the figure 2.4.6c. A similar inverse correlation for Phili and the clear lack of association in Los Angeles and Riverside. . Since the chemistry is likely to be different in the far west, this probably needs further discussion.

Page 2.28, Figure 2.5.1. There must be something wrong here. I t may be in the primary source of the data. If the figure is truly representative of "all age groups" (line 3) then 1.8% time spent in Bar-Restaurant probably is a sampling error. Also if ~60% of people work outside the home how can the Office-Factory be only 5.4%? This seems to me to be an example of uncritical acceptance of data to make a point. I have not tried to evaluate the equations on the next two pages since if the basic data set is wrong there is no point.

Page 2.31, section 2.5.2. Limitations are well described.

Page 2; .35 line18-20. This seems too dogmatic. Clearly at 250m the proximity to sources will affect whether measurements made are an over or underestimate of exposure. Away from sources this might be a well mixed level, which would truly be representative of exposure. Since most urban (and in fact rural) monitoring sites that are not specified as being place for specific emission control are not measuring point sources, the values measured are not likely to be overestimates of exposure for people. Seems sentence on page 2.37, lines 14-15 confirms this.

Page 2.42, Section 2.6 Dosimetry… Almost all reported work is very old. It could have been summarized in much shorter space.

4-6. Integration of evidence on animal tox, human clinical and epi studies sound, balanced and communicated. Is the assessment focused on policy-relevant studies or health findings? Is the discussion of strength, consistency, coherence and plausibility of health effects adequate

The major concern with this chapter (and the next) is that although in places mentioned is made of the potential for studies to underestimate health effects in certain sensitive subpopulations, results (particularly null effects reported) generally do not estimate the effects that are present in the approximately 10-20 and sometimes up to 30 percent of the population being studied who are the true responders (or sensitive subgroups). Their results are often ignored when the summary statistics are given, and yet, they represent a substantial sub-segment of the population who are responding adversely. Throughout these chapters this need to be brought out better. It might be helpful to reverse the order of the presentation and show the human controlled data first. Here it would be easier to bring out the point that there are sub-segments of what are generally believe to be similar healthy people who respond differently. This would point the argument for the epi studies to bring out the sub segment of these populations. The studies discussed are in fact the appropriate ones. The issues of strength, consistency or lack thereof are brought out. However, coherence and certainly plausibility are not well discussed. The end of each section that seems to indicate where the coherence is lacking is often stated as an uncertainty of lack of ability of the writer to come to a conclusion. The raising of the biologic plausibility simple as a statement at the end of a section (see below) is often totally unjustified. SO2 has been around a long time. There is both a large toxicological and human study literature that goes back almost 50 years, without getting into the classified literature that goes back even further. To raise the question of biological plausibility without documenting why the writer thinks such, is simply disingenuous.

Page 3.7, Figure 3.1.1 I do not find this figure very compelling since almost all the estimates are the same (as stated in the text)

Page 3.10, Figure 3.1.3 The graphics in this figure are not clear. The size of the population in each study is reflected in the width of the confidence intervals. But the size of the central tendency estimator (although stated in the description to represent some weighting) is not consistent. The Schwartz study is 300 kids , the Neas 98 yet dots are the same. Also not clear what lag 0-6 means. The last entry is either a sum of all the data and if so not clear the lag used or and inappropriate and really unbelievable value for the 70 children in the Romieu study. Ditto problem with Romieu study in Figure 3.1.4.

Page 3.25, line 14-16. This statement is too definitively negative. The limited data is true, but to conclude that no effect because the tox data was studied at too high a level is premature. This needs to soften to indicate that there are suggestive but limited data. See the way it is said at the end of 3.1.1.4 on airways hyperresponsiveness. The difference between the two sections is 5-10 ppm in animals in the former, and 5ppm in the later.

Page3.30, line 19-21. Here again, the issue seems to be that the animal tox studies simply haven't been done beyond the one species of mice.

Page 3.45, summary of short term exposures. There is not sufficient discussion in this section of the fact that a substantial fraction (15-20%) of the population may be driving the effects noted in normals. Even in those studies in which the overall effects appear to be null there are these "hyperresponders". This may explain why in studies of diseased subjects the positive findings occur more frequently or they appear to be more responsive. Needs more discussion.

Page 3.45 to end of section on Cardiovascular short term effects. Though mention is made of the increased responsiveness in those with preexisting disease, the summary statements in each section seems to ignore this phenomena (in spite of presenting what I would deem as relatively consistent findings that those with preexisting disease are more responsive). Suggest authors look more closely as this phenomena.

Page 3.71 discussion of meta analysis. Not clear if authors are planning to redo this analysis. The suggestion is that the wrong data were used. Is this the comment made by the original author or by the writer of this document? In either case need to clarify what was used, and indicate its limitations or change it.

Page 3.80 and section of short term mortality. The last line of the summary line 5-6 seems inappropriate. The purpose of this section was to comment on the epidemiology. That done to throw in "absence of strong biological plausibility" shows bias rather than good sense. The previous 80 pages in the chapter provide several arguments that there is a biologic reason for looking at SO2 as a potential putative pollutant that affects biological systems. Enough said, so that I do not insult the author further.

7. Public Health Impact
Page 4.3, line 10-11. This is a repeat of an above theme. The Gong study (among others) clearly identifies a sensitive sub-group even among a larger group that would all be considered potentially sensitive. The public health impact is clear that we need to quantify better who these particularly sensitive groups are. This issue comes up several time in the chapter. Although the tables define the likely groups with regard to respiratory conditions, it seems not complete. Perhaps a couple of more summary tables with different parameters are necessary (E.g. Age, sex, other preexisting disease). I would think that this might be useful if population estimates are to be made in subsequent risk assessment documents.

Page 5.2, line 29. I would have thought that a gradient should be going up. Therefore should this be an "west-to-east" gradient rather than as written?
Section 5.2.2.1—Agree with conclusions on short term peak exposures
Section 5.2..2.2—Agree with all except Inflammation. This is too firm a negative as the issue really hasn't been studied effectively.
Section 5.2.2.3—Short term cardiovascular effects—Stated as "inconclusive" is too conservative. Need to bring into this section more that subjects or patients with preexisting disease appear to be more responsive. This would make the conclusion more consistent with a subsequent section (5.2.2.5) short term effects on mortality, since much of the mortality reported is likely to be cardiovascular.
Section 5.2.2.4—other systems, agree.
Section 5.2.2.6—Long term effects on morbidity. Agree
Section 5.2.2.7—Long term effects on mortality. Do not agree. The epi argument is fine. The lack of a plausible toxicological mechanism, is because of a lack of study rather than the actual testing of any mechanisms.
5.2.2.8 Concentration-Response function—There is a distinction here that needs to be made. Although there is a lack of data for a threshold the evidence that there is a concentration-

response function (above some undefined threshold level) is rather good. In most of the studies there is either an increase in response with increasing dose or a greater recruitment of "sensitive subjects" with increasing dose.

5.2.29 Sensitive subgroups. This might be expanded somewhat. Although I would agree with groups specified, I still have concern that within groups there are individuals who are more or less sensitive. What this means is that for those subgroups who are deemed "inconclusively responsive" or with "weak evidence" of responsiveness there are individuals who are truly and unequivocally responsive but we are just not smart enough to subdivide the groups adequately to detect them and define them as a specific subgroup. I guess this is why the original framers of the Clean Air Act thought it was a good idea to have "a margin of safety".

Comments from Dr. Wyzga

Chapter 2

Figure 2.4-1. Since SO2 emissions will be reduced greatly in the near future, can the estimates of future emissions be given in this figure?

P. 2-36, l 10: since the slopes are negative, it would be preferable to use a word other than predictor; e.g., significantly associated with.

Chapter 3

There are several studies that are not included in this chapter. Some of them are quite recent and may not have appeared after the draft of the ISA was written; others report negative results; hence "SO2" may not have been included among the list of key words in literature searches. A comprehensive review of the literature should include them, however. I cite the following:
Ho et al. (2007) Environmental Research 104:402-409.
Ko et al. (2007) Thorax: 62:779-784.
Ko et al. (2007) Clinical and Experimental Allergy 37:1312-1319.
Sinclair and Tolsma, (2004) J. Air & Waste Manage. Assoc. 54:1212-1218
Klemm et al, (2004), Inhal. Toxicol. 16: Supplement I, 131-141
Metzger et al. (2007) Epidemiology 18(5):585-592
Peel at al (2007) Am J Epidemiology, 165(6):625-633

p. 3-6, l. 12: replace "likely due to" with "but there was".
l. 22: insert at end "for CO and NO2, but not for PM10" if Figure 3.1-1 is correct.

p. 3-42: it would be helpful to indicate the levels of SO2 in the various studies; it could provide some clue as to why results are divergent.

p. 3-59: Why is Metzger et al. not included in this figure?

p. 3-59: The conundrum is that the central eastern cities had higher levels of SO2.

Chapter 4

P. 4-4: Studies are cited which used the default convergence criteria with GAM. This is clearly stated, but the implications of this use should be noted.

p. 4-6, l. 8: It should also be noted that exposure measurement error can interfere with estimation of thresholds.

Comments from Dr. Postlethwait

1. In reading the document, especially immediately after the NOx ISA, it would appear import to address whether the epidemiologic cohorts overlapped in geographic locale and/or time with other studies which identified differing pollutants as "causal" to deleterious health outcomes. In other words, were the same people counted more than once? Because of the apparent potential confounding from other pollutants in numerous studies, as noted within the ISA, such an analysis might prove useful in selecting which studies provide the most robust support for direct causality of SOx induced health effects.

2. It appeared that a number of the cited studies reported SO_2 concentrations within ranges that Chapter 2 identified as potentially problematic in terms of measurement error. Consequently, there is either a modest disconnect between stated measurement concerns and the health outcomes assessments or the ISA is simply accepting reported concentrations at face value. The ISA would benefit from this being clarified.

3. This reader found some of the concluding statements among Chapters 3-5 to lack internal consistency with regard to effects levels between clinical and epidemiological studies. In general, it appears that ambient concentrations are approximately an order of magnitude (or more) less than levels that induce observable pathophysiologic effects during controlled studies. Thus, while controlled studies show reproducible effects at $\geq$ 500 ppb, such levels are rarely attained during environmental exposures. Consequently, it would be useful to include, if possible, what short term peaks are attained within the US to provide a stronger basis for causal and biological plausibility statements laid out in Chapter 5. This becomes especially important when one considers the dosimetry data that suggests little SO_2 penetrates to the distal lung and how reported outcomes correlate with estimates of intrapulmonary distribution.

4. There appeared to be select suggestions regarding what types of additional studies are needed and when biological plausibility appeared to be evident. A more consistent format regarding additional studies suggestions and assessment of the mode/mechanisms of biological plausibility would strengthen the document.

5. In general, it appeared that the overarching take home message was at times suggesting that the majority of reported health outcomes were either substantially confounded due to co-pollutants or the lack of clearly observable effects. However, in other portions of the ISA there were fairly dogmatic statements regarding robust associations and biological plausibility. The document would be improved by minimizing such divergent conclusions.

6. Because of the recognized co-pollutant confounding, the very low ambient levels, and a relative paucity of mechanistic data, it is suggested that a short section be included in the end that identified key areas of research needed to reduce the levels of uncertainty.

Comments from Dr. Gordon

Charge Questions

Question 4. The discussion and integration of evidence of sulfur dioxide health effects is quite good in this ISA. Unlike the recent first draft of the NOx ISA, the health effects chapter is of the appropriate length and efficiently discusses the key studies and health effects in a manner that leads to the obvious conclusion that respiratory hospital admission and ER visits, as well as acute bronchoconstriction in asthmatics, are the adverse health effects of concern for sulfur dioxide. The integration of the human clinical and epidemiology studies in the ISA, however, was much better than the animal toxicology data. Admittedly, there are sparse animal toxicology data at concentrations relevant to ambient exposure levels, but Chapter 3 should be restricted to include animal exposure studies of 1 (possibly 2) ppm sulfur dioxide or less. Such levels of exposure are appropriate for interpretation of and integration with epidemiology studies which evaluate associations generally below 50 ppb or with 1 hr peak values higher than 100 or 200 ppb. The high exposure levels used in the described extra-pulmonary studies (e.g., nervous system effects) lend little to our understanding of the mechanisms by which sulfur dioxide causes adverse respiratory or extra-pulmonary effects in susceptible people. For communication purposes, it's important to have an integrated analysis that draws key conclusions from the available epidemiology and toxicology data sets and includes the magnitude of the concentration response for the different health endpoints – this latter is the key to Chapters 3 and 5 but also the overall quality of the ISA.

Question 6. While the ISA has drawn important conclusions regarding the robustness of the respiratory epidemiology data and a lack robustness in the cardiac mortality data, the issue of confounding co-pollutants deserves more attention and discussion. This is critical to the use of the ISA in the Health Plan's risk assessment because the data suggest some studies on sulfur dioxide's role in adverse respiratory effects stands up to inclusion of single co-pollutants while others do not. Given that the high concentrations of sulfur dioxide used in many of the reported animal toxicology studies make plausibility interpretations quite difficult, it is even more critical to clearly discuss the epidemiology findings given the potential confounding by co-pollutants. Along that line, the discussion and integration of the adverse effects of multiple pollutants would be aided by a more complete discussion of numerous studies, particularly animal toxicology and human clinical studies, which have investigated the interaction of particles with sulfur dioxide. These studies are presented in the Annex Tables, but the ISA only mentions one human clinical study (Koenig, 1983) whereas there are several animal studies that demonstrate that particles, especially in the present of moisture, can enhance the effects of sulfur dioxide. These studies may not help in setting benchmark values for health effects, but they will provide biologic plausibility for the benchmark health effects which occur at very low sulfur dioxide concentrations.

Question 7. The respiratory effects (hospital admissions, ER visits, acute bronchoconstriction) have been correctly identified, discussed, and justified in the ISA.

<u>Major Comments</u>:

As key chapters, Chapters 3, 4, and 5 are fairly well-balanced in integrating the exposure and health data needed for risk assessment. The chapters are of appropriate length and detail and do integrate the science of SOX's health effects without providing too much detail (a repeated exception would be the inclusion of animal studies utilizing high ($\geq$ 5 ppm) of sulfur dioxide). Chapter 4 has fairly similar goals to other parts of the ISA (summarizing the adverse health effects and discussing which concentrations and time frames are of concern), and EPA should consider combining the chapters or more clearly delineating why topics are included in each chapter. For example, a large part of Chapter 4 relates to a discussion of susceptible subpopulations, information which could be included in Chapter 3 amongst the other health effects discussions on atopics, asthmatics, and children/elderly.

The bronchoconstrictive response of asthmatics to sulfur dioxide is fairly rapid. EPA should expand their data collection to include 15 min (or shorter) interval data for use by epidemiologists to examine associations between shorter interval peak exposure levels and adverse pulmonary endpoints in future studies.

Medium Comments:

Throughout the ISA, the animal toxicology studies should be included only when they are of relevant concentrations.

Minor Comments:

Chapter 2

page 2-3, line 31 – Stating that sulfur is bound to amino acids implies somewhat that amino acids are grabbing or binding free sulfur rather than that sulfur is a key component of some essential amino acids.

page 2-12 – An additional figure or table which presents the decreasing ambient concentration of sulfur dioxide over the last 2 decades would be helpful. Although the text states the decline in sulfur dioxide concentration over time, a clear graph showing trends in 1-hr, 24-hr, and annual values would be more valuable than 2-yr trend example data for sulfur dioxide and sulfates in a few different cities.

page 2-13 – Figure 2.4-5 needs a better explanation of what is being presented. This same comment applies to all figures in the ISA – the figure legend should allow the reader to fully understand the figure without searching through the chapter's text.

page 2-13, line 1 – The use of 'aggregate' is unclear. The use of 1-hr data to estimate daily or annual values is important but what does 'aggregate up' mean?

Chapter 3

page 3-4, line 1 – Sulfur dioxide concentrations units are usually given as ppm, yet there is inconsistent use of µg/m3 included sporadically throughout this chapter – ppm values should be sufficient.

page 3-10 – In the Figure legend, the 'size of the box of the central estimate' doesn't really add much information when there is already a mean (95% CI) – it seems redundant with 95% CI. Also, when the 95% CI is very large, the box becomes so small that it is hard to see the dash representing the OR.

page 3-15, lines 22-26 – In the ISA, it is a good decision to not present every study that finds the same result, but such sections would be more lucid if information was presented on why some

studies deserve detail and others don't (i.e., better studies or equal in quality but the same findings?).
page 3-15, line 29 – Were the increases of 15 and 23% due to sulfur dioxide alone or sulfur dioxide plus NaCl?
page 3-16, line 5 – typo – delete first 'of'
page 3-19, lines 2-6 – Something is unclear in this section. If the pooled data from mild and moderate/severe groups totaled 40, why does line 6 state 15/40 'moderate/severe subjects'?
page 3-19, lines 8 – 23 – This is an important point/study in understanding whether asthma severity is linked to responsiveness to sulfur dioxide. It is somewhat unclear as the first sentence states that the moderate/severe asthmatics had the greatest physiological and symptomatic responses, then the paragraph points out various sides of the interpretation.
page 3-20, line 9 – typo? Change 22 to 24.
page 3-24, lines 1-13 – Why is this animal toxicology data included by itself when there is data going back to the 1980's investigating vagal pathways in human subjects exposed to sulfur dioxide. Two examples:

Myers DJ, Bigby BG, Calvayrac P, Sheppard D, Boushey HA. Interaction of cromolyn and a muscarinic antagonist in inhibiting bronchial reactivity to sulfur dioxide and to eucapnic hyperpnea alone. Am Rev Respir Dis. 1986 Jun;133(6):1154-8.

Yildirim Z, Kilic T, Koksal N, Kotuk M. Protective effect of ipratropium bromide on bronchoconstriction induced by sulfur dioxide exposure during apricot sufurization processes that causes asthma-like syndrome in agricultural environment. Pharmacol Res. 2005 May;51(5):479-82.

page 3-25, lines 4-13 – These animal studies using high concentrations of sulfur dioxide have little relevance to the ISA.
page 3-26, line 16 – Unclear. Cut 'using Mch …relatively'?
page 3-27, last para - These animal studies using high concentrations of sulfur dioxide have little relevance to the ISA.
page 3-28, line 15 –16 – Should this read 'sensitized to Ascaris' or challenged with Ascaris in already sensitized sheep?
page 3-30, lines 24-25 – Pneumonia and bronchitis (acute) are not usually included collectively under COPD, are they? Later, in Figure 3.2-3, data is given separately for COPD and pneumonia.
page 3-32 – Figure 3.1-7 – Again, the size of the box interferes with interpretation of the OR and 95%CI. A couple boxes are so big that they obscure the reader from seeing if the 95%CI are above the 1.0 Relative Risk level.
page 3-34, lines 15-16 – Is it appropriate to give a risk factor per 10 ppb, when the range given for sulfur dioxide annual means is only 0.9 to 4.8 ppb?
page 3-35, line 13 – ibid
page 3-36, Figure 3.1-9 – Typo? 25 degrees C. The legend says 2 different risk ranges (10 and 40 ppb) – should the differences be delineated in the figure?
page 3-37, Define NR in the figure
page 3-38, line 3 – Age should be stated for these results.

page 3-40, lines 13-17 – This conclusion statement should possibly be qualified a bit more. The sulfur dioxide RR does not appear to be robust in the Schwartz (1995) and Thompson (2001) studies which have fairly tight RR/95% CI.
page 3-41- Define * in the figure.
page 3-45, lines 29-30 – HRV is predictive for survival after an MI, but is there strong evidence for this statement as written?
page 3-48, line 24 – These animal studies using high concentrations of sulfur dioxide have little relevance to the ISA (nor is it referenced).
page 3-54 – The figure has some shading that is not explained. Also, the ED visits and hospital admissions data in this figure are not 'clearly distinguished' as stated on page 3-53, line 8.
page 3-60- This page describes high dose animal studies and should be summarized in a sentence or two or cut out completely. Also, the final para has no exposure concentration or ref.
page 3-61 – Given the high doses used in the animal studies, Table 3.1-1 should be eliminated. The first para on this page may have high concentrations of sulfur dioxide (not given) and could be cut as could the Singh study in the last line of this page (32 ppm sulfur dioxide).
page 3-62 – ibid
page 3-81, lines 1-2 – This is not a chronic study and is a high concentration (10 ppm).
page 3-84, line 17 – 'respiratory health' would include Lung Function (next section), so this statement is a bit too broad.
page 3-84, lines 1-2– This is not a chronic study and is a high concentration (5 ppm).
page 3-87, line 17 – typo? 65/72?
page 3-93, line 16 – Unclear – are the mg/day personal doses?
page 3-103, lines 5-6 – Are risk estimates of 1.02 and 1.04 really different (i.e., 'smaller' as stated)?

Chapter 4

page 4-1, lines 30-32 – The sentence needs editing.
page 4-5, line 6 – It is puzzling why the Ponce de Leon study described here, thus giving it weight/importance but it is one of many in a figure in Chapter 3 and only appears as a reference in Chapter 3.
page 4-9, lines 13-16 – This sentence is speculative. Why wouldn't irritative effects be seen in atopic children?
page 4-9, lines 30-31 – Typo? Repeated on next page.
page 4-10, lines 2-8 – Is the averaging time available for this paragraph?
page 4-10, line 32 – Typo? line ends abruptly.
page 4-12, line 26 – The inclusion of the Ponce de Leon study in the statement regarding the association between sulfur dioxide and hospital admissions in children contradicts the statement on page 3-34, line 18 stating other European studies did not find a significant association.
page 4-14, line 26 – Typo? 'expression' or?
page 4-15, line 31 – Define TNF-1 allele – homozygous for the variant or wildtype?
page 4-18, line 8 – These percentages differ from those given on lines 1-2 on the next page.

Chapter 5

page 5-2, lines 18-24 – While the accurate monitoring of any pollutant is essential to understanding its health effects, is there really a need to push the detection limit even lower? Health effects appear to occur significantly above 1 ppb.
page 5-3, lies 14-31 – This is a very good section, but it would be improved if some additional indoor/outdoor ratios ranges were given. The stated range of 0.03 to 1.01 implies this broadness is frequently the case (the dosimetry section and table in Chapter 3 says otherwise).
page 5-8, lines 1-10 – This section and the supporting studies presented in Chapter 3 provide 'clear evidence for sulfur dioxide effects with peak exposure.' Therefore, it is not clear why a peak exposure (5-15 min) Tier III evaluation was ruled out in the Health Assessment Plan document.
page 5-9, line 22 – Typo? 'symptoms' under Airway hyperresponsiveness subsection?
page 5-9, line 32 – Typo – among = 'amount'
page 5-14, lines 9-14 – This is one of a few possible explanations why population thresholds may be obscured. More importantly, why would this be labeled 'obscured'? A sensitive subpopulation would bring up the lower end of the curve and they are truly responders. I'm not sure the shape of the line matters here as much as identifying the concentration that causes an adverse effect.
page 5-16, lines 14-20 – Do these measurement uncertainties really complicate 'our ability to attribute' if the measurement levels are significantly below observable effect levels?
page 5-16, line 25 – This statement is incorrect – human clinical studies do examine sensitive subpopulations such as asthmatics.
page 5A-2 and 5A-3 – These are all high dose studies and should be cut.
page 5A-2 – define NR; Under the Mortimer study, it states '0-75 ppb (shown in graph)' – what graph?

Comments from Dr. Thurston

General Comments

This document is an excellent first effort at reviewing and integrating the large body of toxicological, clinical and epidemiological knowledge regarding the human health effects of sulfur oxides (SOx). I especially like the evidence evaluation approach outlined on pages 5- and 5-7. It makes a compelling case for the need for a new short-term SO_2 standard (of 1 hour or less averaging time) in order to sufficiently protect public health. However, the document has some shortcomings in several regards, as follows.

First, although the document states on pages 1-1 and 1-2 that "the possible influence of other atmospheric pollutants on the interpretation of the role of SO2 in health effects studies is considered, including interactions of SO2 with other pollutants that co-occur in the environment (e.g., nitrogen oxides, carbon monoxide [CO], ozone [O3], particulate matter [PM])", this is not yet sufficiently accomplished in this document. In particular, the document does not yet sufficiently address the interactions of SOx and PM, though much of the information necessary to do so is already included in the document. As discussed at our July meeting, this document should comprehensively consider the SOx-PM interaction, considering the potential for the potentiation of each pollutant by the co-presence of the other. Yet this is not drawn out in the document, and the PM is only really considered as a possible confounder (i.e., a problem in discerning SOx associations) in this document, rather than as a potential facilitator, of SOx effects. For example, in the dosimetry section, no mention is made (on 2-43) of Dr. Mary Amdur's work showing that the respiratory effects of SOx on animals are greatly enhanced by the co-presence of particles, presumably because particles can absorb the gas and act as an "vector" for the SOx, allowing it to bypass absorption in the upper airways. (One biologically plausible mechanism might be that the adsorption of SOx could be making the metals in particles more acidic and more bioavailable, and thus the SOx makes the particles more toxic.) Similarly, the clinical studies and epidemiology need to consider the potential for a SOx-PM interaction more consistently.

Throughout the document there is some evidence consistent with the PM interaction influence on SOx toxicity, but it is not really considered collectively: one example with regard to clinical studies appears at the top of page 3-16, where it says "One human clinical study provided evidence that during exercise, peak exposures (10 min) to SO2 at concentrations of as low as 0.5 ppm in the presence of hygroscopic particles that can carry SO2 deeper into the lung can elicit significant changes in pulmonary function in asthmatic adolescents. ". Also, with regard to the London Fog Episode on page 3-63, it says: "the 1982 AQCD could not resolve the relative roles of these two pollutants and suggested that the clearest mortality associations were seen when both pollutants were at high levels ". Thus, this interaction is touched upon here and there, but needs to be organized and brought together, and thereby considered in a more "holistic" way. Indeed, these issues need to be handled comprehensively in *both* the SOx and NOx documents Overall, while there are smatterings of references (here and there) to PM-SOx interactions as an possible "confounder" in various passages, I see PM as the insufficiently addressed "elephant in the room" of each of these two new gaseous pollutant assessment documents. Interestingly, this factor is relied upon in making conclusions in Chapter 5 (at the bottom of pages 5-16 and 5-17), but the support for this important point is not well enough developed in the prior chapters. I recommend that the NO_x and SOx documents both address this gas-particle interaction issue more directly and comprehensively.

Second, although the introduction to the document states about Chapter 3 (on page1-5) that " The focus of this chapter is on the strength of underlying epidemiological or toxicological evidence and the coherence and plausibility of the body of evidence for effects on the respiratory, cardiovascular, or other system.", these criteria for the evaluation of the health effects considered are not sufficiently set out at the start of Chapter 3, nor are these criteria consistently applied throughout the document. Presumably these choices are based upon Sir A.B. Hill's 1965 treatise on causality, but it should be explicitly referenced, and the rationale for the selection of these specific criteria from Hill's longer list, and how they will be applied, needs discussion here. I feel that a more consistent application of the A.B. Hill criteria across the various sections, especially as a function of pollutant averaging time and concentration when possible, would enhance the value and usefulness of the SOx document.

Third, and related to the above discussions, the use of epidemiological results with multiple pollutants considered simultaneously is not useful, yet it is done throughout the epidemiology sections. The problem is that including correlated pollutants simultaneously makes the individual pollutants' coefficient estimates biased and largely uninterpretable. As stated on page 35 of the EPA's companion SOx Scope and Methods document: "When collinearity exists, inclusion of multiple pollutants in models often produces unstable and statistically insignificant effect estimates for both SO_2 and the co-pollutants." Thus, one should consider at most two pollutants at a time, and even then only as a sensitivity analyses, not as useful estimates of the individual pollutants effect estimates or their significances. In multi-pollutant models, only the linear combination of the pollutant effects is an unbiased estimate (not the individual coefficients), so a more useful approach might be to look at the total effect of all pollutants in such models, and see if that overall estimate is increased by the addition of another pollutant in order to evaluate if some additional information is provided or not. At a minimum, the epidemiology tables should remove all models considering more than two pollutants at a time. In addition, the interpretation of multi-pollutant models should consider the potential for pollutant interactions (e.g., possible potentiation of effects) as possible explanations for variations across models, studies, and locales, rather than merely dismissing pollutant terms affected by the inclusion of other pollutants as indicative of statistical confounding. The effects of co-pollutants on the SOx term may reflect real biological interactions of effects, especially between PM and SOx, and this should be considered as a possibility, in addition to statistical confounding.

Fourth, units are not yet consistent throughout the document. Sometimes it refers to ppm, sometimes to ppb, and sometimes even ug/m^3 (e.g., see top of page 3-12) . Personally, I prefer ppb, but some consistent concentration metric should be used throughout to ease cross-discipline and cross-study comparisons.

Specific Comments

Chapter 1

Page 1-1, lines 23,24: This sentence is muddled, seeming to include particulate sulfates among gaseous SOx. Clarify.

Pg. 1-3, line 8: Add a comma after "possible".

Pg. 1-3, line31: Change "systems" to "capabilities". The U.S. and Canadian health care systems are not the same.

Page 1-4, last paragraph. Fix grammar. Each consideration is written as a question, but there is no question mark. I suggest not having them as questions, changing each numbered consideration to start "whether" instead of "were" or "are", and fixing each from there.

Chapter 2.

Pg. 2-3, line 18. Note that nearly 90% of the power plant emissions of SOx are from coal fired power plants.

Pg. 2-8. line 29. Note that, as shown in Figure 2.4-1, most of this improvement in SOx emissions was made prior to 1995, and that progress has slowed in the last decade.

Pg. 2-13. Also present the distribution in the cumulative frequency of occurrences of 5-minute average SO_2 concentrations exceeding given levels (e.g., %>50 ppb, %>100 ppb, %>250 ppb, etc.) in the U.S. data collected in the past 10 years (1997-2006) at some 300-400 ambient sites, as discussed in the draft SOx Scope and methods for Exposure and Risk Assessment document (Nov. 2007).

Pg. 2-24, last par.: Published evidence documenting intercontinental transport should be added. For example, from China to the Eastern U.S., see: Lall, R and Thurston G. Identifying and quantifying transported vs. local sources of New York City $PM_{2.5}$ fine particulate matter air pollution. *Atmospheric Environment* 40 (2006) S333–S346.

Pg. 2-43. last par.: Add discussion of Dr. M. Amdur's work showing the copresence of particles increases penetration and effects of SOx. This will provide support needed for statements made on page 5-16, line 30.

Chapter 3.

Page 3-1: Add discussion of A.B. Hill's Criteria for causality and lay out the criteria to be used to evaluate the various studies and to evaluate causality. This should be consistent with the system utilized in Chapter 5 (pg. 5-6), and be applied and considered throughout Chapter 3. If a section or a study doesn't help evaluate one of these criteria, eliminate it. Add studies only if they are needed to address a criteria not yet considered in each section of this chapter.

Page 3-2, line 10: Change "have little ability" to "cannot definitively"

Page 3-2, lines 14-15. Change the words "and serves as an important tool in addressing the issue of confounding by copollutants." to read: ", and may provide some insights into the potential for confounding or interactions among pollutants."

Pg. 3-2, line 26. Change "multiple model" to read: "multiple pollutant models

Pg. 3-3, line 1. Start new par with: "While clinical…", and drop "do in fact"

Pg. 3-3, line 17. Add summary sentence noting that, given their respective strengths and limitations, all the different types of toxicological, clinical, and epidemiological studies are needed in order to evaluate the causality of SOx-health effects associations with confidence.

Pg. 3-16, lines 4-7. This study is quite important, given the discussions on pages 5-16 and 5-17. This needs further discussion as to how it fits with toxicological studies that have shown greater effects with the co-presence of particles (e.g., Amdur, et al.), and how this shows coherence across disciplines. It may also account for why epidemiology studies show effects at lower levels than controlled SO_2 studies, as PM is always also present in the ambient environment.

Pg 3-19, line 11, change "response with" to "response to"

Pg. 3-23, 2nd par. Note the Koenig study showing effects at lower levels with PM co-exposure, and highlight effect of exercise. Both of these factors important to understanding the epidemiology, and why it might see associations at lower levels than most controlled pure SO2 studies.

Pg 3-26, lines 10,11: Use ppb, not ug/m^3 .

Pg. 3-35, lines 3,4,5. Eliminate sentence. Such many pollutant models are uninterpretable, and should not be cited.
Pg. 3-40, line 4. Expand this paragraph to consider biological interactions between SO2 and PM that may be causing the multi-pollutant model results, not just statistical confounding.
Page 3-40. lines 29-32. Note that the warmer months are generally associated with higher activity levels by children, and more acidic forms of SOx, so these epi results are consistent with the toxicology and clinical results.
Page 3-44, line 2. What averaging time for the 40 ppb? How about for the other values in this paragraph?
Pg. 3-47, line 17. Beta blockers not a factor in this study?
Pg. 3-56, line 9. Use ppb, not ug/m^3 .
Pg. 3-58, line 5. Consider interactions in this paragraph, too, especially for PM. There is a need for evaluation of coefficients as a function of PM levels, not just whether the coefficients change in simultaneous multi-pollutant models. More advanced and comprehensive approaches are needed to sort this out than considered here.
Pg. 3-58, line 19. Change "likely confounded" to "may be confounded"
Pg. 3-59, Table3.1-13. Replace all models considering more than two pollutants at a time with two pollutant models when available. Are PM, SO2 two pollutant models available for most studies? If so, consider making a table of SO2 alone vs. SO2 with PM.
Pg. 3-67, line 9. change "confounded by" to "confounded with".
Pg. 3-74, line 9. Change title and text to include both confounding and interactions. Note where epi evidence supports or refutes and interaction between sulfur dioxide and PM (e.g., where single pollutant model SO2 effect larger in places with higher PM).
Pg 3-75, lines 17-18. Change to read: "….and Europe generally suggest that SO2 mortality risk estimates may be confounded by co-pollutants, making a definitive distribution of effects among the pollutants difficult."
Pg. 3-80, line 4. Add comma after "is suggestive"
Pg. 3-80, line 5. Add discussion of possible interactions between SO2 and PM,a nd evidence for or against within the epi literature. New tables and/or analyses of the results from each study may be needed to add this perspective on the literature.
Pg. 3-87. Add discussion of the ACS study (pope et al, 2002) vis-à-vis SO_2 and lung cancer mortality.
Pg. 3-96, line 23. Elaborate on the fact that this lack of specificity by SO_2 undermines its credibility as causal for long-term mortality in this study. PM is specific, SO_2 is not.
Pg. 3-101, line 23. Good point. Also, note that the Harvard 6-Cities study has a lower % College Educated (closer to the overall U.S.), and gets higher PM RR estimate, consistent with this conclusion.
Chapter 4.
Pg. 4-3, line 11. Note Koenig study showing effects at lower levels in co-presence of PM, as in the real world case. Also, note that exercise lowers threshold for effects.
Pg. 4-5. Note that the co-presence of PM and its varying interactions with SO_2 may also affect ability to detect a threshold in epi studies.
Pg. 4-7, line 22. Note also that children are active, and exercise lowers the threshold for SO_2 effects in clinical studies, so that is another way in which children are more at risk.
Pg. 4-10, lines 21-26. Rewrite with respect to the criteria to be outlined in the beginning of Chapter 3, based upon A.B. Hill's criteria for causality.

Pg. 4-16, line 10. Change “only very limited” to “insufficient”.

Chapter 5.

Page 5-1, line 23. Change “coal and oil” to “coal”

Page 5-1, line 27 from “substantially since 1990” to “substantially since 1990, but progress has slowed in the last decade”.

Pg. 5-14, line 16. Note the increased susceptibility resulting from exercise and co-presence of particles in this section.

Pg. 5-15, line 12-13. Change “only very limited” to “insufficient”.

Pg. 5-15, line 24. How high is >120? Give actual value.

Pg. 5-16, lines 28-31. Good discussion, but this needs better documenting in the body of the document (e.g., Chapters 2 and 3).

Pg. 5-17, lines 16-21. Good discussion, but needs better documenting in the body of the document. Especially for the effects of SOX on particle bioavailability of transition metals.

Page 5-18. line 1. Change “will result in decreased” to “will be associated with a decrease in the”

Comments from Dr. Hattis

Integrated Science Assessment

2. Spaciotemporal patterns and exposures, policy-relevant background.

I think the document could have gone a little farther in analyzing the data in Table 2.4.2 on SO2 concentration distributions observed by existing monitors in CSMA's for different averaging times. Figure 1 shows lognormal plots of the data in this table. From the correspondence of the data points to the fitted straight lines, it can be seen that particularly for the shorter averaging times, the data are well described by lognormal distributions. In the fitted regression line the intercept is an estimate of the logarithm (base 10) of the geometric mean and the slope is an estimate of the logarithm of the geometric standard deviation. For example, the estimated geometric mean for the maximum 1 hour daily averages of the readings from CSMA monitors is 100.806 = 6.4 ppb and the estimated geometric standard deviation is 100.524 --about 3.34. These results allow us to make at least some quantitative estimates of the likely frequency of ambient outdoor exposures at levels associated with various incidences of short term responses to SO2 in populations that have been studied in clinical settings. I will use these results for my response to the charge question (#7) on likely public health impacts below.

4. Integration of evidence on health effects from toxicological human clinical and epi studies technically sound, balanced, and clear?

5. Integration of health evidence focus on the most policy-relevant studies and health findings?

6. Conclusions drawn on the strength, consistency, coherence and plausibility of health effects.

Generally I agree with the somewhat skeptical treatment of the epidemiological results as likely to be confounded with effects of particles. I think this is particularly likely because although we know how to measure SO2 gas pretty well, we don't know exactly what the most causally related components of particles really are. Because SO2 and fine particle levels are correlated, this makes it quite likely that existing regression related findings are attributing some of the mortality and other responses causally related to some particle fractions (size distributions, composition) erroneously to SO2.

Particularly in the light of this, I would have preferred a more quantitative treatment of the issue of human variability in the undoubted causally related responses observed from clinical exposures to SO2. I am particularly intrigued by the possibility of a more quantitative analysis of the individual subject response data of Horstman et al. (1986) reproduced in Figure 3.1-6, and any other similar data sets.

For the analysis of human variability in Figure 2 below I have extracted the individual Horstman et al. data as best I could from the figure provided in the ISA and the accessible abstract (I could not easily obtain the original paper). Figure 2 is based on the conventional assumption for probit analysis that in the population of asthmatics studied there is a lognormal

distribution of individual thresholds for the response (a doubling of airway resistance during exercise). In this case the intercept is an estimate of the log of the SO2 level needed to elicit the response in the median asthmatic (100.0189 = 1.044 ppm = 1044 ppb) and the slope is an estimate of the log of the geometric standard deviation [the Log(GSD) in our terminology] of individual response thresholds (100.374 = 2.37).

In previous efforts my colleagues and I have compiled a substantial database of information on human interindividual variability for a variety of responses (see the website at http://www2.clarku.edu/faculty/dhattis). The log(GSD) of about 0.37 in this case is not at all unusually large-it is actually toward the lower end of observations of variability in responses to acute inhalation exposures compiled in our data base (Table 1) (however, it can be seen that in many of these cases with larger variability the agents act via specific receptors or via allergic processes that may well in general be subject to more variability than responses to nonspecific irritants).

Given the variability analysis in Figure 2, it is straightforward to make at least a tentative projection of the likely incidence of responses for asthmatics similar to those studied by Horstman et al. (1986) at any air level, assuming that the population distribution of response thresholds is in fact perfectly lognormal:

ppb	expected incidence of response (% of days expected to cause 100% increase in specific air way resistance for exercising asthmatics, ignoring the exposure duration difference between 10 minute studied exposure and 1 hour duration for the greatest 1 hour average in a 24 hour period)
10	3.4E-08
20	2.2E-06
30	1.9E-05
40	7.7E-05
50	2.1E-04
100	3.2E-03
150	0.012
200	0.028
400	0.13
600	0.26
800	0.38
1000	0.48

7. Public health impact and characterization of susceptible/vulnerable groups.

Given the analyses presented earlier in my responses to charge questions 2, and 4-6, and

* Assuming that both the exposure distribution and the distribution of individual thresholds for response in asthmatics are perfectly lognormal,

* Ignoring for now the exposure duration and intake difference between the one-hour exposures measured by the monitors and the 10 minute exposures used to measure effects in the exercising asthmatics studied by Horstman et al., and

* Neglecting any systematic differences there are likely to be between individual personal exposures and air concentrations measured in the elevated outdoor compliance monitors

we can derive an estimate of the overall fraction of days that asthmatics similar to those in the studied group. We do this by cutting the assumed lognormal distribution of air concentrations from 0 to 1000 ppb into intervals of 1 ppb, calculating the number of asthmatic people who might be in each interval, and summing up the number likely to respond during the maximum hour's exposure on each day (Table 2). Overall the fraction of asthmatic-days expected to elicit a response of the severity recorded by Horstman et al. (1986) is about 2.9 per 10,000. Interestingly, half of the total response incidence is attributable to very rare high exposures (over about 230 ppb). This results from the larger estimate of variability in exposures, compared to the estimate of variability in human response thresholds.

8. Adequacy of first draft to provide support for future risk, exposure and policy assessments?

There are useful data here that can support future risk, exposure and policy assessments. As illustrated by the tables and figures I have constructed for the responses to previous charge questions, I think the current document can be taken further in quantitative analysis of exposure levels, and the extent of variability in individual humans' susceptibilities/distribution of response thresholds. I also think that the authors should make some more ambitious attempt to draw quantitative conclusions via meta-analytic combination of the very best multipolutant studies of effects. This will inevitably still be confounded by the fact that we don't have a good idea what the appropriate causally relevant metric is for the particle exposures (in terms of size fractions and composition). Absent this, the fact that it is much simpler to measure SO2 levels, and the fact that SO2 and fine particle levels are correlated means that multiple regression analyses are very likely to attribute some of the effect caused by particles to SO2. Getting a quantitative handle on how large this confounding really is will be very challenging with existing information.

Figure 1

Lognormal Plots of Data from Table 2.4.2--Distributions of S02 Concentrations (ppb) for Different Averaging Times

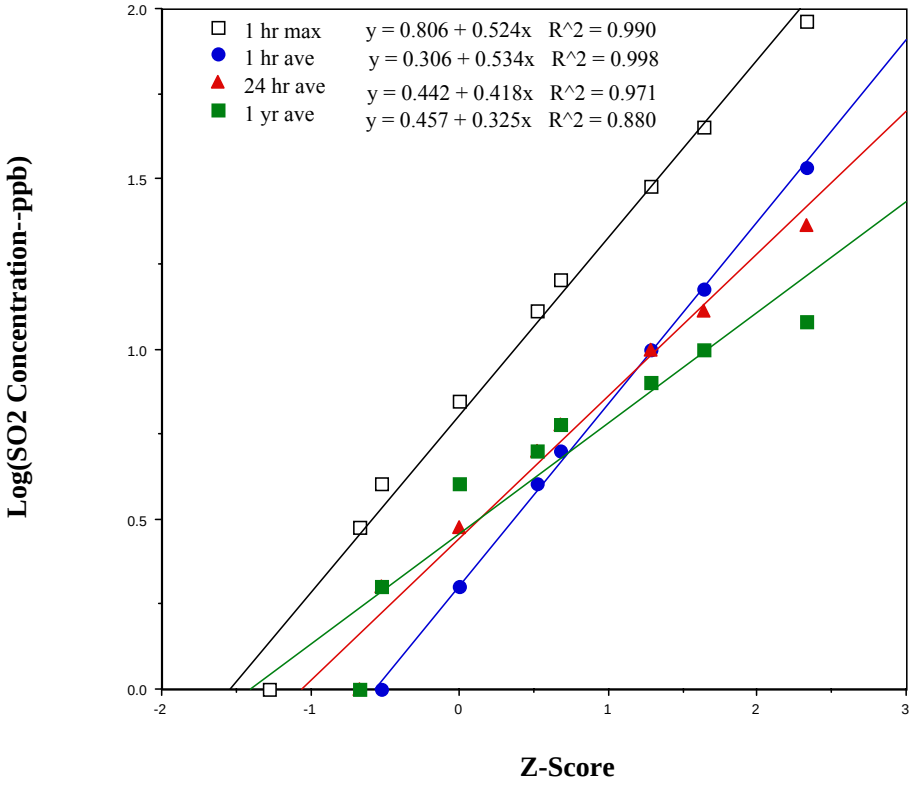

Figure 2

Lognormal Plot of the Distribution of Individual Sensititivities (SO2 Concentrations Needed to Double Specific Airway Resistance) For 27 Exercising Asthmatics (Horstman et al. 1986)

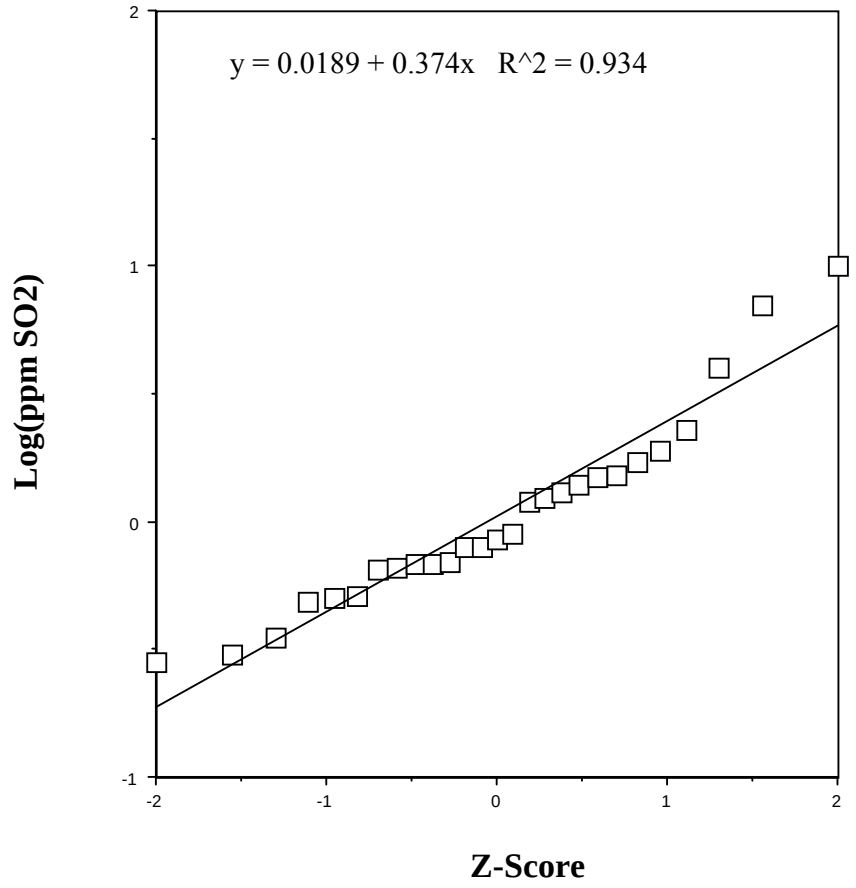

Table 1
Previous Observations of Human Interindividual Variability in Local Lung Function Responses to Inhaled Agents

log(GSD)	response studied	population studied	N	agent	data source
0.74	Air Conc. Needed to cause 10%, 15%, and 20% decrease in FEV1	Females--general population	748	Methacholine	Paoletti et al., 1995
1.00	Air Conc. Needed to cause 10%, 15%, and 20% decrease in FEV1	Males--general population	810	Methacholine	Paoletti et al., 1995
0.32	FEV1 change in relation to CXT of ozone exposure (clinical)	Experimental subjects		Ozone	McDonnell et al. 1995 analyzed in Hattis 1998
0.43	FEV1 Increase by Antiasthmatic	Asthmatics	14	Salbutamol	Lipworth 1992, analyzed in Hattis 1998
0.76	PD20--concentration needed for 20% increase in individual baseline value of FEV1	Atopic subjects	13	Ragweed allergen	Meerschaert, 1999
0.57	PD20--concentration needed for 20% increase in individual baseline value of FEV1	Atopic subjects	17	Histamine	Meerschaert, 1999
1.33	Specific Airway Resistance PC50--concentration needed for 50% increase in individual baseline value	Bakers--occupationally exposed	34	Wheat flour dust	Merget, 1997
1.11	Specific Airway Resistance PC50--concentration needed for 50% increase in individual baseline value	Bakers--occupationally exposed	34	Wheat flour extract	Merget, 1997
0.64	Specific Airway Resistence--concentration needed for 20% increase in individual baseline value	5733 smokers with mild to moderate airflow obstruction	5733	Methacholine	Tashkin et al., 1996
0.42	Specific Airway Resistence--concentration needed for 100% increase in individual baseline value	Healthy athletic adults, 18-50	66	Methacholine	Balmes et al., 1997
0.51	Specific Airway Resistence--concentration needed for 15% increase in individual baseline value, mean of 2 trials with and without ozone	Allergic asthmatic patients	9	Grass allergen	Hanania, 1998
0.78	Specific Airway Resistence--concentration needed for 15% increase in individual baseline value, mean of 2 trials with and without ozone	Allergic asthmatic patients	6	Ragweed allergen	Hanania, 1998
1.13	Specific Airway Resistence--concentration needed for 20% increase in individual baseline value	9 year old New Zealand Children	813	Methacholine	Sears et al., 1996
0.60	Specific Airway Resistence--concentration needed for 20% increase in individual baseline value	Allergic asthmatic patients	15	Methacholine	Hanania, 1998
0.97	Specific Airway Resistence--concentration needed for 20% increase in individual baseline value	General adult population, Norwegian community, Age 18-73	490	Methacholine	Bakke, 1991
0.59	Specific Airway Resistence--concentration needed for 20% increase in individual baseline value	Nonsmoking adults with mild asthma	17	Histamine	Evans, 1996
0.27	Specific Airway Resistence--concentration needed for 20% increase in individual baseline value	Nonsmoking adults with mild asthma	18	Metabisulphite	Evans, 1996

Source: Human interindividual variability database, updated as of 5/05, available "http://www2.clarku.edu/faculty/dhattis" discussed in

Hattis, D. "Distributional Analyses for Children's Inhalation Risk Assessments." Journal of Toxicology and Environmental Health, 71:1-9, 2008 in press.
Hattis, D. and Lynch, M. K. "Empirically Observed Distributions of Pharmacokinetic and Pharmacodynamic Variability in Humans-Implications for the Derivation of Single Point Component Uncertainty Factors Providing Equivalent Protection as Existing RfDs." In Toxicokinetics in Risk Assessment, J. C. Lipscomb and E. V. Ohanian, eds., Informa Healthcare USA, Inc., 2007, pp. 69-93.
Hattis, D., Baird, S., and Goble, R. "A Straw Man Proposal for a Quantitative Definition of the RfD," Drug and Chemical Toxicology, Vol. 25, pp. 403-436, (2002).

Table 2
Illustrative Calculation of the Expected Fraction of Days on Which Exercising Asthmatics Might Experience a Doubling of Specific Airway Resistance, Subject to Extensive Assumptions (see text)

Upper end of conctration interval (ppb)	overall contribution to fraction of days with response in interval	cumulative total fraction of days with response
10	1.6E-09	1.6E-09
20	9.0E-08	9.1E-08
30	5.4E-07	6.3E-07
40	1.4E-06	2.1E-06
50	2.6E-06	4.7E-06
60	3.9E-06	8.6E-06
70	5.1E-06	1.4E-05
80	6.2E-06	2.0E-05
90	7.0E-06	2.7E-05
100	7.7E-06	3.5E-05
110	8.2E-06	4.3E-05
120	8.5E-06	5.1E-05
130	8.7E-06	6.0E-05
140	8.8E-06	6.9E-05
150	8.8E-06	7.8E-05
160	8.7E-06	8.6E-05
170	8.6E-06	9.5E-05
180	8.4E-06	1.0E-04
190	8.2E-06	1.1E-04
200	7.9E-06	1.2E-04
300	6.4E-05	1.8E-04
400	4.0E-05	2.2E-04
500	2.4E-05	2.5E-04
600	1.5E-05	2.6E-04
700	9.7E-06	2.7E-04
800	6.4E-06	2.8E-04
900	4.3E-06	2.8E-04
1000	3.0E-06	2.9E-04

Comments from Dr. Samet

General Comments

The draft ISA for SO_x follows the model used for the NO_x ISA and consequently many of the general concerns that I expressed about the earlier ISA are applicable to the SO_x ISA. In my general comments on the NO_x ISA, I noted the failure to adequately describe the methodology for the review, the lack of transparent criteria for evidence evaluation, and the incomplete development of a causal framework for interpreting the findings of epidemiological studies. Interpretation of regression estimates for "SOx effects" is particularly problematic. This and other problems noted for the NO_x ISA are again evident and equally limiting; I attach my earlier comments on the NO_x ISA.

On reading this ISA, I was particularly concerned by the failure to explicitly describe the approach for evidence evaluation from the outset. The authors' conclusions reached on the various health outcomes are summarized in the attached table; these were largely expressed in summary paragraphs at the end of subsections. The depth of analysis was limited and terminology is not uniform. Toxicological information appears to have been variably considered in making judgments on the strength of evidence.

These evident limitations of the ISA appear to reflect inadequate development of the overall approach to preparing the ISAs. In response to a request for a protocol for carrying out the ISAs, I was informed that there was not a formal protocol beyond the brief procedural charges transmitted by Marcus Peacock in memoranda dated December 7, 2006 and April 17, 2007. In my opinion, a systematic review process should not be implemented, absent a formal and documented protocol that describes the approach for evidence gathering, evaluation, summarization, and interpretation, including uniform criteria and language for describing the strength of the available evidence. This deficiency of the process needs immediate discussion.

The Chapter 5 summary of the evidence and judgment as to the strength of evidence for causation is not well grounded in the review offered in Chapter 3. I concur with the judgment as to the causal nature of the short-term effect on lung function, derived largely from experimental findings in human clinical studies. The ISA uses the terms "consistent and robust" in referring to the findings on respiratory health, phrases that are not well supported by the judgments made in Chapter 3 (see summary table).

Comments on Chapter 5 and responses to charge questions 5 and 6

General Comments on Chapter 5: Chapter 5 presents a bulleted summary of the findings of previous chapters, noting through this display, the advances in evidence since the prior reviews of SO_x. The approach to providing the updates is succinct, although the methodology for determining these advances lies on the more opaque approaches of the prior chapters. Of concern is the methodology for integrating the health findings, which as noted, is set out only briefly in Chapter 5. The focus is primarily on the "positive" findings in a large and difficult body of evidence. While criteria such as coherence and consistency are mentioned, I note a failure to be truly integrated. Rather, the focus immediately settles on the findings of the human exposure studies with regard to short-term effects of experimental exposure on lung function, and selected positive findings in regard to respiratory health. The latter do not receive adequate interpretation.

Charge Question 5: Potentially, the clinical studies are policy relevant in certain settings, specifically those where short-term exposures comparable to the concentrations used in the exposure studies might occur. The groups at risk would include exercising persons as well as those with asthma, particularly if exercising. This is likely a relatively infrequent exposure scenario (to be addressed in the risk assessment). The ISA does not adequately consider the potential concentration-response relationships for this short-term outcome. Given understanding of the dosimetry of SO_x in the respiratory tract, what doses would be anticipated at short-term peaks likely to be experienced at present? The respiratory health outcomes are policy relevant and much of the literature relates to exposures that would be encountered in the general population setting. In this regard, the findings are more relevant to the present NAAQS.

Charge Question 6: The conclusions drawn in the draft ISA are incompletely grounded in considerations around strength, consistency, coherence, and plausibility. The draft ISA is neither sufficiently comprehensive nor thoughtful in its application of these criteria. The criterion of strength refers to the magnitude of the effect. With regard to current levels of ambient SO_x, "strong" effects would not be anticipated and "weak" effects are far more plausible. The document has no explicit criteria for consistency and plausibility is variably addressed in the various integrative sections. As I note above, criteria for evidence evaluation are offered, only briefly, in Chapter 5 and there appears to be little consistency in Chapter 3 in reaching judgments.

Response to Charge Question 8: Based on my review of this first external draft ISA, I find it to be deficient, perhaps largely reflecting an inadequately developed synthesis methodology. I am doubtful that the document will become adequate without development of a more explicit protocol and rigid adherence to it.

Table Summary Conclusions Concerning Various Health Outcomes in the SOx ISA

INDICATOR	GROUP	FINDING
Respiratory Outcomes/Short-Term		
Respiratory Symptoms	Children	Associated
Respiratory Symptoms	Adults	Mixed
Lung Function	Children	Mixed/no independent effect
Lung Function	Adults	Effects in clinical studies of peaks
AHR and Allergy	Adults/Children	Suggestive of increase in AHR
Host Defenses	Adults/Children	Weakly suggestive epi findings
Emergency Medical Care for respiratory outcomes	Adults/Children	Suggestive evidence for association
Cardiovascular Outcomes/Short-Term		
CVD/Short-term Exposure	Adults/HR and HRV	Some suggestive findings
	Adults/Repolarization Changes	No Conclusion
	Adults/Arrhythmias	Inconsistent
	Adults/BP	No effect
	Adults/Blood Markers	No effect
	Adults/Acute MI	No evidence for increased risk
	Adults/CVD Emergency care	Collective evidence weak for association
	Adults/Cardiac Emergency Care	Weak
	Adults/Stroke Emergency Care	Inconsistent
Mortality	Overall	Positive coefficients but possible confounding; overall, evidence is suggestive but limited
	Cardiovascular/Respiratory	Association but possible confounding
Respiratory Morbidity/Long-Term		
	Respiratory Health	Evidence is suggestive, inconsistencies
	Lung Function	No indication of an effect
	Carcinogenesis	Unlikely to have an effect
Reproductive outcomes	Children	Difficult to draw conclusions
Long-Term Mortality	Adults	Several studies suggestive, limited interpretation for causality

Specific Comments

Page: 1-3

Number: 1 Author: JSAMET Subject: Sticky Note Date: 11/26/2007 3:01:51 PM
Confusion here in terminology and concept. Synergism is one form of effect modification and interaction and effect modification are often used interchangeably.

Page: 2-38

Number: 1 Author: JSAMET Subject: Sticky Note Date: 11/26/2007 3:58:52 PM
But certainly, outdoor-indoor relationships are probably highly variable within and across communities.

Page: 3-6

Number: 1 Author: JSAMET Subject: Sticky Note Date: 11/26/2007 4:16:13 PM
Or, the effect of SO2 occurs through secondary PM formation

Page: 3-40

Number: 1 Author: JSAMET Subject: Sticky Note Date: 11/26/2007 4:32:57 PM
This is a rather "loose" way to summarize a complicated body of evidence. This is not a sufficiently transparent finding.
1

Page: 3-42

Number: 1 Author: JSAMET Subject: Sticky Note Date: 11/26/2007 4:35:12 PM
Again, not clear what is the basis for this summary judgement of a very mixed body of literature.

Page: 3-44

Number: 1 Author: JSAMET Subject: Sticky Note Date: 11/28/2007 12:43:07 PM
Not clear how epi studies contribute to biological plausibility.

Page: 3-74

Number: 1 Author: JSAMET Subject: Sticky Note Date: 11/26/2007 4:47:25 PM
This discussion refers to confounding based on changes in estimates comparing single-pollutant with multi-pollutant estimates. Such changes may have other explanations related to mediation of effects by secondary PM and differing degrees of measurement error for correlated pollutants.

Page: 3-101

Number: 1 Author: JSAMET Subject: Sticky Note Date: 11/26/2007 4:58:14 PM
What is meant by "concentrated? Main point is that the high exposures are largely in the East?

Page: 5-1

Number: 1 Author: JSAMET Subject: Sticky Note Date: 11/26/2007 6:12:33 PM
This sentence is conceptually vague and needs to be expanded into an introductory paragraph that is far clearer.

Page: 5-6

Number: 1 Author: JSAMET Subject: Sticky Note Date: 11/28/2007 3:20:42 PM
These would appear to be the "rules of evidence" for interpretation of the findings. They need description and development in Chapter 1. Was this set of criteria given to all auithors? Was the same approach used in the NOx ISA?

Comments from Dr. Ultman

Chapter 1. Well-written chapter. The inclusion of the framing questions, as was done in the recent ISA for ozone is a good idea. On lines 8-9 of page 1-2 where current SOx standard is given, I would add values in $\mu g/m^3$ as well as ppm.

Chapter 2. This chapter rather seamlessly synthesizes a diversity of material from source-to-dose. With respect to the dosimetry section, there was no mention of the extrapolation of animal exposure conditions to equivalent exposure concentrations in humans using mathematical modeling or default scaling methods. It is not clear whether this was an oversight, or if there is nothing in the literature specific to SOx that adequately informs such an animal-to-human extrapolation. Also absent from the dosimetry section was the notion that (because reaction products of inhaled SO_2 can reach the bronchial and pulmonary circulations) exposure of the respiratory system to SO_2 might lead to downstream systemic effects.

Some specific suggestions and editorial comments for the authors to consider follows:

Page	Lines	
2-1	29	Define what the symbols OH and M represent, and show OH as a free radial.
2-2	2-2	You may want to show HO_2 as a free radical.
2-9	Fig 2.4-1	Change styles of bar fills so that legend entries are more obvious, or use color plates in final ISA.
2-10	Fig 2.4-2	Change to grayscale instead of pseudocolor in final plots.
2-11	Fig 2.4-3	Change to grayscale instead of pseudocolor in final plots.
2-13	1-2	It's not clear how this sentence follows from the previous sentence.
2-13	Fig 2.5-5	Contrary to the explanation in the text and the figure title, this is not a box plot. Also, it would be helpful to add another plot of this type showing the distribution among all cities of the daily-average concentrations, and how this varies throughout the year.
2-14	15	Pearson's correlation coefficient between what two variables?
2-26	Fig 2.4-11	Change to grayscale instead of pseudocolor in final plots.
2-29	10	This equation assumes that ambient and non-ambient exposures are additive. This is not necessarily the case. It is more accurate to say that concentrations arising from ambient and non-ambient sources are additive as represented in Eq. 2-11.
2-29	26	Equation 2-11 is really a "recasting" of equation 2-10. Is there a final equation for E_T that is missing?
2-43	20	Nodelman and Ultman did not make any measurements of SO_2.
2-44	6-8	It was not "total respiratory absorption" but rather "absorption efficiency that was probably independent of inspired concentration. From this point to the end of the chapters the authors must be more careful to specify whether which of these two quantities they mean.

Chapter 3. This is a well-organized chapter covering numerous important issues that arise in the identification of health effects and sensitive populations.

Page	Line	
3-8	Fig 3.1-2	Why does the the right-most portion of the solid curve break up into isolated points? extrapolated from data at lower concentrations?
3-9	25	At this point and throughout the chapter, the phrase "lower respiratory symptoms" is frequently used. "Lower respiratory tract symptoms" would be a more accurate phrase.
3-13	30	Add "of" after "7/8."

Chapter 4. An important point in this chapter is that public health impact should be evaluated in terms of the magnitude of a health effect as well as the number of people it affects. This is particularly important when trying to identify susceptible subpopulations.

Chapter 3 contained several tables illustrating the odds ratio or relative risk associated with SOx exposure and how this varied in different subpopulations. On the other hand, chapter 4 contained tables and text that documented the size of these subpopulations. Yet, no where in chapter 4 are the two synthesized so that the reader can appreciate the possible health impacts.

Chapter 5. This chapter would be more effective if it was organized around the framing questions laid out in chapter 1. It would then be more apparent whether or not the ISA exercise was successful. As it stands now, it is difficult and, in some cases, not possible to infer the answers to the framing questions from the conclusions made in chapter 5.

Comments from Dr. Balmes

12-4-07

Charge 4 To what extent is the discussion and integration of evidence from the animal toxicology and controlled human exposure studies and epidemiologic studies technically sound, appropriately balanced, and clearly communicated?

GENERAL COMMENTS

Like its counterpart in the NOx ISA, Chapter 3 on Health Effects is long (104 pages) and overly detailed in certain parts. There should be less detail about specific studies in the chapter text; these details are best left to the annex. That said, Chapter 3 in the SOx ISA does a better job of summarizing and evaluating the evidence in each section than the previous effort. Still, by trimming detail and endeavoring to present the information in a more thematically clear manner (i.e., each section should have a clear "story line"), a revised chapter will better support whatever recommendations for an air quality standard emerge from the review process. The chapter as currently written reads too much like a mini-criteria document rather than an integrated synthesis.

In general, the presentation of the results of the animal toxicological, controlled human exposure, and epidemiological studies that have been reviewed is technically sound. However, the criteria for selection of specific studies in all three categories should be clearly stated. In addition, the criteria for judging the strength of findings from specific studies as well as those used to assess aggregate findings of studies on a relevant research question should also be clearly stated.

I am concerned about one section (3.1.1.4 Airway Hyperresponsiveness and Allergy). On pp.3-27 and 3-28, there are several paragraphs that discuss results of animal toxicological studies in support of the statement that "A limited number of animal studies also suggest acute SO2-induced increases in airway obstruction in allergen-sensitized guinea pigs and sheep." This statement is the lead sentence for a paragraph that discusses three papers (Douglas et al., 1994; Lewis and Kichner, 1984; and Scanlon et al., 1987). First, none of these studies involved allergen-sensitized guinea pigs or sheep. Second, only one of these studies (Douglas et al., 1994) involved allergen-sensitized animals. Third, none of these studies provide evidence in support of the statement. The next paragraph in this section discusses a paper by Riedel et al. (1988) in a manner that misrepresents the results of the study. The Riedel paper actually presents results that show that SO2 exposure enhances sensitization to allergen in guinea pigs rather than "the effect of SO2 exposure in ovalbumin-sensitized guinea pigs." The study animals were exposed to ovalbumin only on the last 3 days of a 5-day exposure protocol; one week after the end of exposure to various concentrations of SO2 or filtered air the animals were challenged to ovalbumin and the SO2-exposed animals, at all concentrations, were more sensitive to ovalbumin than the control animals. The next paragraph opens with the statement "Similar findings were observed in which guinea pigs were exposed to a single

SO2 concentration." The Park et al. (2001) paper actually is similar to the Riedel paper because it presents results that suggest that SO2 enhances sensitization, not that "airway obstruction induced by ovalbumin challenge was higher in ovalbumin-sensitized guinea pigs exposed to 0.1 ppm SO2 for 5 days compared to sensitized guinea pigs that werfe not exposed to SO2." The results of the Kitabatake et al. (1992, 1995) and Abraham (1981) papers are correctly discussed. The concluding paragraph of the section contains the statement that "Toxicological studies that observed increased airway obstruction and hypersensitivity in allergen-sensitized animals provide biological plausibility." The Kitabatake and Abraham papers are the only toxicological studies that support the increased airway obstruction part of this statement. The Riedel and Park papers actually do not provide evidence in support of "increased hypersensitivity in allergen-sensitized animals"; rather, these two papers show that SO2 exposure can enhance sensitization to an inhaled allergen. This subsection needs to be rewritten.

In section 3.3.1.2 there should be some discussion of papers by Jedrychowski et al. that present results of the longitudinal study of lung growth in children and adolescents in Cracow, Poland. Although there was exposure to both particulate matter and SO2 in Cracow during the study period, these studies provide evidence of a possible effect of SO2 on growth of lung function in children and merit discussion.

In my view, the epidemiological data are relatively consistent and coherent with regard to the association of short-term exposure to SO2 and emergency department visits/hospitalizations for asthma and all respiratory diseases, particularly among children. The toxicological evidence in Chapter 3 could be more clearly presented to convincingly support potential mechanisms for asthma exacerbation.

The toxicological data that are best presented in the chapter are the controlled human exposure data which indicate that asthmatic individuals are especially sensitive to SO2 exposure in terms of respiratory symptoms and bronchoconstriction. While these data do provide some plausibility for the epidemiological studies that find an association between ambient SO2 and emergency department visits or total hospitalizations for asthma, they do not illuminate how SO2 exposure might induce respiratory hospitalizations in non-asthmatic individuals. The animal toxicological data provide little help because exposures much higher than ambient are required to induce measurable effects. Enhancement by SO2 of airway responses to specific allergen challenge is perhaps the potential mechanism of asthma exacerbation best supported by the animal toxicological data. While these data are discussed in Chapter 3, they are included in a way that somewhat misrepresents their meaning (section 3.3.1.4 and p. 5-9 "Airway Hyperresponsiveness" bullet) in an effort to support the limited epidemiological evidence that SO2 can induces airway hyperresponsiveness in atopic individuals.

In terms of clear communication, Chapter 3 as currently drafted falls somewhat short. The text in Chapter 3 needs to be tighter, less redundant, and more thematically organized (i.e., each section should have a story line). In particular, the summary/integration subsections should provide an overview of the quantity and quality of the evidence for

the health outcome(s) of interest as well as evaluation of how well the toxicological data support the epidemiologic findings.

SPECIFIC COMMENTS

p. 3-3, line 26 The title of section 3.1.1 should make it clear that it is a summary of literature reviewed in previous documents.

p. 3-14, line 19 In epidemiological studies assessing the relationship…

p. 3-16, lines 2-4 …reported in most studies mean that the evidence is insufficient to conclude that short-term exposure to ambient SO2 has an independent effect on lung function.

p. 3-26 and throughout the document Airway hyperresponsiveness (AHR) should be used consistently throughout the document. Refraining from using bronchial hyperresponsiveness (BHR) or airway hyperreactivity (Chapter 4) will avoid unnecessary confusion

p. 3-27, lines 20-21 See comments on section 3.1.1.4 above. This sentence is inappropriate here. The papers discussed in this paragraph do not support this sentence and require a different interpretation.

p. 3-28, lines 19-20 As noted in the general comments above, this statement is only supported by two of the studies reviewed in the previous paragraphs in this subsection (Kitabatake and Abraham).

p. 3-30, line 9 …and no alterations in pulmonary immune system function were reported…

p. 3-32, Figure 3.1-7 legend There is no asterisk in the figure.

p. 3-34, line 3 …and 1-h max in Paris, France.

p. 3-35, line 18 …though not always statistically significant…

p. 3-36, Figure 3.1-9 legend There is no asterisk in the figure.

p. 3-37, Figure 3.1-10 legend There is no asterisk in the figure.

p. 3-39, lines 31-32 In summary, there are limited studies…, making it difficult to draw conclusions…

p. 3-42, lines 8-9 "and the animal toxicological studies that observed SO2 altered lung defenses" should be deleted. The interpretation of the relevant toxicological data on p. 3-30, lines 19-21, is the correct interpretation.

p. 3-45, lines 1-2 Ibid.

p. 3-47, line 20 …and coagulation markers…

p. 3-49, line 8 …NO2, CO, and PM2.5 were found…

p. 3-52, line 28 …but not SO2 or NO2…

p. 3-57, line 4 …the relationship between ambient SO2 and hospitalizations for cardiac disease…

p. 3-59, Figure 3.1-13 legend There is no asterisk in the figure.

p. 3-62, line 30 …no adverse effects on development or reproduction other than the limited evidence of neurodevelopmental effects noted above.

p. 3-63, line 13 Exposure to ≥5-ppm SO2 was found…

p. 3-69, line 11 …for NO2-mortality associations, but were weaker.

p. 3-77, line 14 …CHF 1 year before death were compared.

p. 3-78, line 29 …other constituents…are responsible for the adverse effects.

p. 3-84, section 3.3.1.2 See my comment above about including discussion of results from the Cracow longitudinal study of lung function.

p. 3-88, line 32 …The reliability and validity…have been reviewed…

p. 3-93, line 16 Is this really SO here?

SPECIFIC COMMENTS RE: CHAPTER 4

p. 4-8, line 23 Should be airway hyperresponsiveness here.

p. 4-10, lines 1-2 The last two lines of p. 4-9 are repeated here.

p. 4-12, lines 21-26 The Ponce de Leon et al. (1996) and Anderson et al. (1998) papers should not be cited in support of this statement given the data presented in Figure 4.2-1.

p. 4-14, lines 25-27 …(e.g., homozygosity for the null allele at the GSTM1 and GSTT1 loci, the presence of the Val-105 variant allele at the GSTP1 locus) that significantly affect expression of protein or function in the lung.

p. 4-18, line 8 …with approximately 10% of adults or 13% of children having ever been diagnosed with asthma…

SPECIFIC COMMENTS RE: CHAPTER 5

p. 5-9, lines 18-20 As noted for Chapter 3, the first part of this statement is supported by only two papers cited in the document and the second part is misinterpreted. This bullet on "Airway Hyperresponsiveness" should be rewritten.

Comments from Dr. Pinkerton

Charge 4 To what extent is the discussion and integration of evidence from the animal toxicology and controlled human exposure studies and epidemiologic studies technically sound, appropriately balanced, and clearly communicated?

Charge 5 To what extent does the integration of health evidence focus on the most policy-relevant studies or health findings?

Charge 6 What are the views of the Panel on the conclusions drawn in the draft ISA regarding the strength, consistency, coherence and plausibility of health effects of sulfur oxides?

Charge 7 What are the views of the Panel on the appropriateness of public health impact and the characterization of groups likely to be susceptible or vulnerable to sulfur oxides?

REPLY:

The presentation and relevance of the toxicology, controlled human exposure and epidemiological studies within this report is excellent. The studies are logically interpreted. The length and depth of the material is quite extensive, therefore, some consideration should be given to 1) shorten and/or combine materials from chapters 3 and 4 and 2) include further materials in the Annexes of the ISA.

The authors of the chapter have provided a nice balance for each of the studies referenced and described in the document. The scope of the studies presented is reasonably complete and adequate. The human health studies are a major force to drive the newest and most relevant information to facilitate updating the criteria document for SOX.

A key issue in establishing the next NAAQS standard for SOX is the manner in which SOX is measured and reported. The importance of exposure to SO_2 over the short timeframe of exposure (5 minute peak concentrations) is of high relevance as reflected in much of the writing of the ISA. These short-term analyses are likely to be a better predictor of health effects. Therefore, such measures are critical to be included as part of the analysis in the next consideration of the SOX ISA. The authors provide solid support for the inclusion of such measures which is highly commendable. The importance of inclusion of SOX measures over the short timeframe has been argued previously, but dropped from consideration in the previous review of the document. Therefore, the authors should be applauded for their renewed interest and strong justification for the use of a 5-minute peak concentration in evaluating 1 and 24 hour averaging times. This assessment could provide extremely important insights as the tiered approach

The animal toxicology studies in large measure are performed at SOX concentrations that are orders of magnitude higher than those for human clinical or epidemiological studies. Although these differences in exposure ranges complicate the use of animal toxicology

studies to establish appropriate levels for the protection of human health, never-the-less these studies do provide biologically plausible mechanisms of effect for exposure to SOX.

A critical component of the ISA draft document is to define SOX effects as clearly as possible, while also making it very clear of the ever-present difficulties to deal with confounding co-pollutants. Although it is fairly clear SO2 is the most prevalent form of gaseous SOX, while other forms of sulfur oxides tend to from particulates, it should be done clear SO2 is the critical measure to represent SOX. However, if this is not the case, then further clarification should be made in the document.

A critical review of the most current literature since the last NAAQS document for SOX emphasizes the need for additional research to further explore the importance of genetic susceptibility and whether genetics further influences susceptibility for those in sensitive populations.

There appears to be a general lack of consideration for the sensitivity of newborns and children to the effects of SOX. Asthma has clearly been identified in both young and old that is a factor to increase risk of the effects of exposure to SOX. However, it will also be critical to further investigate the potential and biologic plausibility of children to exposure to SOX that may be different from that observed in adults.

Mortality is clearly an important consideration in assessing the health effects of SOX. However, it is unclear how aggressive will be policy-decision making and/or the factors that must be taken into consideration to factor mortality into a revised criteria document for SOX. Will such a regulation be based on a specific threshold and/or a linear model?

It is also important to note that accurate exposure levels today in the United States today are almost never as high as that level found in the current SOX NAAQS.

Comments from Dr. Sheppard

Overall: The ISA follows a new structure and represents a substantial change from the previous scientific review. EPA is to be commended for a great start in carrying out the intent of the new process.

1. To what extent are the atmospheric chemistry and air quality characterizations clearly communicated, appropriately characterized, and relevant to the review of the primary SO_2 NAAQS? And 2. Are the properties of ambient sulfur oxides appropriately characterized, including policy-relevant background, spatial and temporal patterns, and relationships between ambient sulfur oxides and human exposure?

The characterization needs to more directly address the spatio-temporal variation in SO2 and the contribution of monitor siting characteristics to the summary data because this has direct bearing on the interpretation of the epidemiological studies. Monitors sited near local sources that don't reflect usual population exposure aren't appropriate to be used in time series studies. How prevalent are such monitors? How many cities have only local source monitors?

Analyses should be added to relate 5-minute averages to 1-hour averages to bring forward into the health assessment.

The discussion of correlations is too ambiguous as presented to be useful (p 2-33 and Table 2.5-1).

I don't find the information in the ISA adequate for understanding personal exposure to SO2. Even if measurement data are limited because of LOD issues, there are studies of I/O ratios and (lack of) indoor sources that add important information to the understanding of characteristics of personal SO2 exposure. I would like the ISA to assemble and integrate the research so that summary statements about estimates of attenuation of ambient source concentration for personal/indoor exposure (e.g. α or Finf) are included in the document.

3. Is the information provided on atmospheric sciences and exposure sufficient for the evaluation of human health effects of sulfur oxides in the ISA?

No. See my comments w.r.t. question 2. Make sure to bring forward key points about exposure with respect to evaluation of health effects into chapter 5.

Chapter 2 comments

Presentation in this chapter needs to focus on understanding needed to interpret the health studies and to inform the health assessment.

- Revamp the analysis of AQS data to address key aspects of exposure that are relevant to health studies. This includes adding a thorough analysis of 5-minute averages and their relationship to 1-hour averages, and analyses separating

monitors by proximity to source and other key features of spatial distribution. In order to estimate the susceptible population exposed to SO2, spatial analyses to predict SO2 features by land use characteristics at the level of census tracts is needed. Add in-depth analyses of AQS data to the annex and then bring forward the most important features into chapter 2.

- Include land use and/or point source analyses to help with later analyses that will need to quantify the degree of population exposure to high levels of SO2 from point sources
- Exposure measurement error in time series studies is likely to be a very important factor for SO2 because of the local source nature of this pollutant and the monitor siting policies. (This could actually be an issue for all short-term epidemiological studies, also including case-crossover and panel studies.) Analyses of AQS data to try to clarify this issue are needed. For instance some cities only have SO2 monitors sited near a local source (e.g. Seattle). This has huge implications for estimates from time series studies. Add an analysis that classifies cities according to siting criteria of monitors (e.g. only local source, only population-oriented, both types) and provide results that allow this feature to be incorporated into the evaluation of the epidemiological studies.
- Analyses of personal or indoor concentrations vs. outdoor concentrations need to be summarized in the context of informing the models for personal exposure from ambient and non-ambient source exposures. The summary in the document (section 2.5.3) does not address this adequately. Table 2.5-1 needs to be revamped to clarify the correlations presented. I suggest including scatterplots of daily data and separating individuals/homes in the plots. Analyses need to be conducted so parameters assumed in APEX can be obtained directly from the ISA. (More extensive analyses can be put in the Annex and appropriate summaries brought forward.)

Tables and figures in the chapter need to be thoroughly revamped. Some may be completely revised based on my recommended thorough analysis of AQS and other data that should be added to the Annex.

- The Figure 2.5-2 summary of indoor and outdoor SO2 only shows annual means. This analysis is so highly summarized it is difficult to apply these results to the epi studies. It would be more relevant to look at scatterplots of daily data.
- Figure 2.4-5 is mistitled and mixes data from all monitors regardless of siting criteria. I think it is important to distinguish extremes due to local sources from extremes at locations without local sources because the proportions of the population exposed or not to local sources is relevant to the interpretation.
- Table 2.4-1 only provides means and does not summarize distributions. At a minimum add a measure of spread. Think about how to make this analysis more informative.
- Table 2.4-2 also needs stratification by siting characteristics or other features to be identified from thorough descriptive analysis of the AQS data.
- Table 2.4-3 is too highly summarized to be useful. The summary doesn't inform the interpretation of epidemiological studies. It isn't clear how each of the ratios is calculated. Distinctions such as single vs. multiple sites, single vs multiple

individuals, individual-level vs averaged data, home ventilation characteristics, and presence or absence of indoor sources cannot be discerned from the table. Revamp to allow these results to be used to inform the estimates of Finf and/or α in the personal exposure model.

4. To what extent are the discussion and integration of evidence on the health effects of sulfur oxides from the animal toxicological, human clinical, and epidemiological studies, technically sound, appropriately balanced, and clearly communicated?

There needs to be a protocol for reporting on studies and selecting what information to report. As a strong example of information that should *not* be included, I think it is inappropriate to report the results of finely stratified subgroup analyses. As an example, in summarizing the results of Bozen et al (p 3-27 l 3) the initial sample size of 327 was reduced to a subgroup of 25 for reporting. This result is almost certainly an artifact of the post-hoc stratification of the analysis and reporting. A uniform protocol for selecting papers and results within papers would certainly not allow reporting of such a result.

Chapter 3 comments

Restating a point made with respect to chapter 2, interpretation of all epidemiological studies of SO2 effects (both long-term and short-term) needs to be assessed in the context of exposure assessment, particularly with respect to how evenly distributed SO2 is in space in a certain area, and the effect of local sources, particularly on the monitoring data used in the study.

- I question whether short-term epidemiological studies based on central site monitoring data are reasonable given the local source nature of SO2. Monitor siting needs to be considered as part of the evaluation of the short-term epidemiological studies (e.g. time series, case-crossover, panel) because of the low monitor-to-monitor correlations frequently observed and the prevalence of SO2 monitors sited to capture the effects of local sources. Using a monitor that doesn't represent usual population exposure in these studies is likely to have a fair amount of classical measurement error in the estimates.

It is not adequate to refer to "epidemiological studies" as a synonym for specific epidemiological study designs such as time series studies (e.g. p 5-13; not sure this is an issue in this chapter).

I am concerned that the ISA review and annex summaries for health gloss over very important subtle issues that affect the analysis and interpretation of the studies. Each study has unique strengths and weaknesses. Some features are shared by the design, exposure data, or software, while some are unique to the individual study. While summarizing such a complex set of features will be very difficult to address completely, progress can be made by expanding the detail in the summaries of each study and following a strict protocol for reviewing, summarizing, and evaluating each study. Expanding upon the list of detailed information summarized on page 1-6, the tabular summaries should include (1b) characteristics of the monitoring data used in the study

(2b) study design (2c) summary of analysis approach and key features of the analysis (3b) description of the population selection criteria (5) any major limitations or issues with study, (6) discussion of study strengths, weaknesses, and key features. (I note some of this information is mentioned in individual studies.)

- As an example, here are some features of Schildcrout et al that I would emphasize but which are not included in the ISA or Annex: This paper is stronger than other panel studies of its type because of the additional control for confounding provided by using within-city estimates of the odds ratios and its unique focus on the combined effects of changes in two pollutants simultaneously. The summary should include the average number of observations per subject and unique details about the analysis. Key features of the pollutant data may impact study results, including the extensive imputation of the PM data, the limited availability of the ozone data, and the siting of SO2 monitors in some cities (e.g. the only SO2 monitor in Seattle is right next to a cement plant).

Add comments on interpretation of epidemiological studies. For instance,

- Monitor siting in individual studies is an important feature.
- Discuss multipollutant models and the interpretation of studies where a single effect in the presence of others, i.e., the effect of a change in SO2 with all other pollutants held constant (see e.g. figure 3.1-11 p 3-41) vs. studies where joint effects are reported from changing SO2 and another copollutant simultaneously (see e.g. figure 3.1-1 p 3-7). The role of exposure measurement error could have important ramifications in multipollutant analyses because SO2 may have more measurement error than PM.
- The evidence that the population is heterogeneous with respect to exposure that induces response has important implications for population-level epidemiological studies (e.g. time series studies). Because of averaging, heterogeneity in the threshold is likely to translate into a linear dose-response function in time series studies.

Is it possible to obtain the relevant quantities from epidemiological studies reporting multipollutant model results in order to calculate the joint effects of the change in two pollutants simultaneously? It may be worthwhile to contact authors of key studies and determine if the relevant quantities can be obtained. If so, ISA authors may wish to consider this suggestion a possible comprehensive strategy for reporting epidemiological study results at a given lag:

1. Single pollutant model for SO2 [Independent effect]
2. Joint effects model for SO2 and selected other pollutants (perhaps only PM) (Note: need to think carefully about the assumed increments for comparability with the other analyses) [Mediated effects by another pollutant in addition to SO2]
3. Multipollutant model estimates of SO2 adjusted for PM or selected other pollutants (these others are held constant, again perhaps only PM) [confounding effects]. Similar estimate for the PM (or other pollutant) effect adjusted for SO2.

4. Commentary on interpretation of each of the above effects, along with discussion of the lack of evaluation of effect modification (given the published analyses by and large haven't done this).

7. What are the views of the Panel on the appropriateness of public health impact and the characterization of groups likely to be susceptible or vulnerable to sulfur oxides?

A discussion of responders needs to be added to the public health impact chapter. While such individuals haven't been shown to be a clearly identifiable subgroup they are a significant fraction of the population and deserve attention.

Chapter 4 comments

Variable sensitivity in the population (i.e. presence of responders) is a key issue for regulation and needs to be addressed directly in this chapter.

Address overlap with chapter 3 in revision of this chapter. Look at public health impact from a regulatory point of view and discuss the evidence from this perspective.

5. To what extent does the integration of health evidence focus on the most policy-relevant studies or health findings?

There needs to be a protocol to systematically classify studies with respect to their policy-relevance and use that to weight the evidence. Aspects of this protocol must address adequacy of exposure data, study design, population size and relevance, and publication bias.

6. What are the views of the Panel on the conclusions drawn in the draft ISA regarding the strength, consistency, coherence and plausibility of health effects of sulfur oxides?

The organization of chapter 5 should be reevaluated with respect to the framed policy questions (page 1-2) and information needed in the health assessment. In addition, the criteria for conclusions need to be specified in the chapter, including the classification of health evidence, the selection of studies to focus on, evidence for coherence, etc. The classification of health evidence should be revamped using published criteria from NAS/IOM as a guide.

Chapter 5 comments

Organization: It would be useful to consider organizing this chapter with respect to information that needs to be brought forward into the health assessment. Important questions include are the properties of the atmosphere well measured, what is the relationship between concentration and exposure, what do the monitoring data tell us about the epidemiological studies. In addition or alternatively, the chapter should address and can be organized around the policy questions framed in chapter 1.

Add cross-referencing to previous chapters whenever possible to facilitate the reader's ability to get more in-depth information.

The concluding statement from this chapter (p 5-17 to 5-18) needs to be better substantiated. It needs to be explicitly stated if this statement should be interpreted in the context of the already low usual population exposures to SO2 or regardless of the existing levels.

Section 5.1.3: Need to revise the summary to be more useful to the exposure assessment.

P 5-3 l 19-21: I would like to see a summary statement about the slope or estimate of Finf for SO2.

P 5-3 to 5-4: Add comments about monitor siting and their effect on different study designs, particularly the role this may play in interpretation of time series studies.

P 5-9 to 5-10 summary of respiratory ED and hospitalizations: Here is an example where the alignment of the epidemiological and clinical/toxicological may be helped by comments on the representativeness of exposure to SO2 in time series studies.

P 5-13 l 14-17: Is it justified to base the ISA conclusion so completely on the conclusion reached by the authors of this study?

P 5-13 l 31: Here is a place where the general term "epidemiological" appears to be used for the more specific term of "time series". All users of this document will be better served by the more exact term.

P 5-14 l 4: Include the references here when specific studies are referred to so explicitly.

P 5-14 l 9-10: Here is an example of a place where it would help the reader to cross-reference earlier chapters and/or the annex.

P 5-16 l 14-20: Representativeness of monitors is key for SO2 study interpretation. Futhermore, the ability to separate effects of SO2 from other pollutants is a fundamental one for interpretation of SO2 effects estimated in epidemiological studies.

P 5-17 l 9-10: I would hope that the further analysis of AQS data requested will better inform this statement. If after that analysis the statement is still justified, is there a plausible chemical/atmospheric reason for this?

8. What are the Panel's views on the adequacy of this first external review draft ISA to provide support for future risk, exposure and policy assessments?

Overall, this is a good first start, but as noted in detail above and by other panel members, there are areas that need additional attention or thorough revision.

Additional specific comments

1-5 to 1-6: I note that controlled human studies are not addressed. Is this an oversight?

2-13 Figure 2.4-5 is a scatterplot not a boxplot.

2-14 l 20-21: Correlations are very sensitive to the range of the data.

2-40 l 19-20: A huge population would be needed to get a good estimate of X_t^A, so this is not feasible in practice for any pollutant.

2-40 l 27: The population size is relevant to this argument and should be stated. Also it would be informative to comment on how results would compare for SO2 if they were available.

2-41 l 1-7: A comment on the applicability to SO2 is needed here since it is so highly reactive.

2-41 l 11: Replace "epidemiological" with "time series".

2-41 l 12-14: It is more correct to say the time series studies estimate a different parameter because they use concentration instead of exposure.

2-42 l 7-16: This paragraph needs work. Attenuation of an effect estimate could indeed change conclusions. The complex soup of pollutants that exist in the atmosphere could play an important role in health but the chapter hasn't addressed this directly. Is this the place to discuss incorporating evidence from clinical and toxicological studies?

3-26 l 10: Insert the number of tests per person, e.g. "range of __ to __ per person". Here is an example where deeper understanding of the key features of the study may help with interpretation.

3-84 l 21: The cross-sectional design is a weak design so I would not put much weight on this result.

3-84 l 28: Does "no consistent" mean "not significant"? Elaboration about consistency should *not* be a discussion of positive vs. null findings with no point estimates or confidence intervals reported.

4-1 l 30 to 4-2 l 1: Other possible reasons are heterogeneity of attenuation of ambient concentration and exposure measurement error due to monitor siting near local sources.

4-3 l 5-8: With only 3 dose levels, it is difficult to fit much other than a linear function.

4-3 l 19-26: Has the Lin et al study been evaluated to assure readers there has been adequate control for seasonal confounding? There is limited mention of this in the annex, but the case-control design suggests the extra attention is merited.

4-3 l 27: Please not the design, not just the general term "epidemiological".

4-4 l 5-6: I think there are many features of a study that are more important to mention in the ISA than the GAM convergence criteria.

4-5 l 15-25: Can this discussion be made clearer? Wasn't some of the work done by these authors (Brauer et al) published in the peer-reviewed literature? See for instance Brauer, Brumm, Vedal, Petkau 2002 in Risk Analysis, Vol 22, p 1183-93.

5-3 l 19-21: It would be more useful if this bullet were to focus on a summary of the slope of the association and not merely a description of the strength of the association.

5-3 section 5.1.3: Add a bullet on the effect of monitor siting on exposure, particularly with respect to different study designs.

Appendix Table 5A-3: Is it worth adding additional percentiles so there is some hope that more studies will have some air quality data represent