UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON, D.C. 20460

OFFICE OF THE ADMINISTRATOR
SCIENCE ADVISORY BOARD

August 22, 2008

EPA-CASAC-08-019

The Honorable Stephen L. Johnson
Administrator
U.S. Environmental Protection Agency
1200 Pennsylvania Avenue, N.W.
Washington, D.C. 20460

Subject: Clean Air Scientific Advisory Committee's (CASAC) Peer Review of EPA's *Risk and Exposure Assessment to Support the Review of the SO2 Primary National Ambient Air Quality Standards* (First Draft, July 2008)

Dear Administrator Johnson:

The Clean Air Scientific Advisory Committee (CASAC), augmented by subject-matter-experts to form the CASAC Sulfur Oxides Primary NAAQS Review Panel conducted its review of EPA's *Risk and Exposure Assessment to Support the Review of the SO_2 Primary National Ambient Air Quality Standards*: First Draft (REA) on July 30-31, 2008. The REA was prepared by the EPA Office of Air Quality Planning and Standards (OAQPS) staff as part of EPA's ongoing review of the primary national ambient air quality standards (NAAQS) for sulfur dioxide (SO_2). The CASAC held a subsequent teleconference on August 12, 2008, to discuss its draft advisory letter. This letter presents CASAC's advice on the first draft REA.

The CASAC found the draft provides an appropriate framework for a risk and exposure assessment. However, it is incomplete and lacks certain details as discussed in the Panel's answers to the Agency charge questions that follow. In addition, a major concern is that the benchmark values being considered for SO_2 effects are based on the benchmark values (0.4 to 0.6 ppm) suggested by the *Integrated Science Assessment (ISA) for Sulfur Oxides – Health Criteria* (Second External Review Draft, May 2008) to which CASAC raised objections in our letter to you dated August 8, 2008. The CASAC believes strongly that the weight of clinical and epidemiology evidence indicates there are detectable clinically relevant health effects in sensitive subpopulations down to a level at least as low as 0.2 ppm SO_2. These sensitive subpopulations represent a substantial segment of the at-risk population. The benchmark values should be adjusted downward in both the ISA and the REA documents. A second major concern is that the overall approach to risk characterization remains to be specified.

CASAC's response to EPA's charge questions are summarized below. Individual recommendations from CASAC Panel members to strengthen the final REA are appended in Enclosure B.

Air Quality Information and Analyses:

Charge Question 1: We have evaluated SO_2 air quality throughout the United States, using all available 5-minute and 1-hour ambient monitoring data for years 1997 through 2007. To what extent are the air quality characterizations and analyses technically sound, clearly communicated, appropriately characterized, and relevant to the review of the primary SO_2 NAAQS?

Chapter 6, at present, provides a good overview of existing air quality data, with particular focus on the limited 5-minute average data and how to extend the more widely available 1-hr average observations. At present, the chapter is often difficult to follow, and the need for specific analyses is not always readily apparent. In the beginning, the chapter should provide a road map for what is done and why, and the material in the chapter should be better focused. In regard to the analyses, EPA should consider describing the observations using log normal distributions. Several figures need to be clarified as recommended in the individual comments.

Charge Question 2: To what extent are the properties of ambient SO_2 appropriately characterized, including ambient levels, spatial and temporal patterns, relationships between various averaging times, and the relationship between ambient SO_2 and human exposure?

EPA staff uses the limited data available for 5-min peak concentrations by developing a model of 5-min peak concentration to the related 1-hr average concentration, then applies that model to the more widely available 1-hr data to generate estimated peak concentrations. At present, the approach chosen, in part, uses calculated coefficients of variation (COVs) to classify monitors, and develops cumulative density functions (CDFs) for the peak-to-mean ratios from which to apply Monte Carlo simulations. EPA should assess whether more general parametric relationships (possibly assuming log-normal distributions) between the 5-min and 1-hr average concentrations can be derived, and thus minimize the occurrence or extent of differences between the predicted and observed numbers of exceedances. For those cases where very large differences occur, more careful analysis is needed. In addition to addressing outliers at the high end of the distribution, more attention to the distribution at lower levels is warranted, particularly considering the CASAC's recommendation to consider potential exposures below 0.4 ppm. The current analysis is confusing because of changes in the monitoring network that affect raw temporal trends and inference about the relationship of ambient concentration to human exposure. EPA should develop analytical approaches that facilitate intended interpretations, such as presentation of concentration trends not biased by network modifications.

Charge Question 3: Twenty locations were selected for detailed analyses, using ambient SO_2 monitoring data for years 2002-2006. What are the views of the panel regarding the appropriateness of these locations, the time period of analysis, and the approach used to select them?

While the choice of the first three locations is very solid, the REA needs to better justify its choice of the remaining 17 locations for detailed analysis. Restricting the analysis to counties where 3 or more monitors exist may exclude locations where peak concentrations are most likely. Although the 3-monitors-per-county criterion is reasonable, other criteria are equally justified (e.g., assessing where peak concentrations are most likely to occur).

Charge Question 4: In order to simulate just meeting either the current 24-hour or annual standards, staff adjusted SO_2 air quality levels for the years 2002-2006 upwards in all but one location. Ambient monitoring data in North Hampton County PA were above the 24-hour standard in the year 2006 and were therefore adjusted downward. To what extent is the approach taken technically sound, clearly communicated, and appropriately characterized?

The approach used is reasonable, although it could be more clearly communicated.* While the panel recognizes that this approach assumes, somewhat unrealistically, that all source emissions would increase similarly, this simple approach is adequate for the purposes of the REA.

Charge Question 5: What are the views of the Panel regarding the adequacy of the assessment of uncertainty and variability?

The uncertainty and variability discussion, at present, is mostly qualitative and needs work. It should be more quantitative in regards to the development and use of peak-to-mean ratios (PMRs) with particular attention to the uncertainty in the model's ability to simulate exceedances over a wider range of levels. Further, the influence of monitor siting (e.g., terrain, source location, magnitude of emissions) and "upset" emission conditions on measured concentrations and distributions needs to be more thoroughly considered. The discussion of ambient monitor to exposure representation recognizes the many limitations of how observations may represent actual exposures, but should be more definitive as to the degree of uncertainty and possible bias that is introduced. With regard to the assessment of uncertainty introduced by the statistical model (Section 6.5.7), the REA should be explicit as to the use of duplicate or independent data, and should consider the use of data-withholding approaches.

Exposure Analysis:

Charge Question 1: To what extent is the assessment, interpretation, and presentation of the initial results of the exposure analysis technically sound, clearly communicated, and appropriately characterized?

The presentation in Chapters 2 and 7 should be revised to better frame the exposure analysis. The committee suggests the text describe how the different components of the exposure assessment fit together, with perhaps a schematic diagram displaying the steps of calculation. The two intended applications of the exposure predictions should be discussed along

* The REA should be explicit in stating that the analysis was limited to the primary standards and did not consider the three-hour secondary SO_2 NAAQS.

with an overview of the features of the simulated exposure data that are most important for each particular application. There is currently no discussion of the health risk assessment application of the exposure modeling or the key features of the exposure predictions that will influence the risk assessment.

While the exposure analysis was generally technically sound, there were several specific concerns. The estimation of the additional eleven 5-minute concentrations in an hour forces all the other values near the hourly average. This artificially reduces the variability of the 5-minute data and effectively constrains the predicted number of 5-minute exceedances in any given hour to be one or less. Furthermore, given the inertia effect of the indoor air diluting the peak levels, the temporal correlation between outdoor 5-minute levels may be critical to correctly calculating the distribution of 5-minute average levels indoors. In addition, the assumed removal rate distributions may strongly affect conclusions since they are represented as uniform and with high rates. The data underlying these removal rate distributions must be fully described, in addition to adoption of reasonable methodology to derive distributional inputs for this factor for use in current modeling approaches such as APEX. Because the indicated ranges will lead to very large reductions in expected SO_2 indoor concentrations, this is a key issue for the modeling of indoor exposures and it is likely to be a large source of uncertainty.

Charge Question 2: The draft risk and exposure assessment evaluates exposures in selected locations encompassing a variety of SO_2 emission source types in the state of Missouri; these areas were chosen as an initial case study since 1) air quality measurements indicated numerous exceedances of 5-minute benchmark values, 2) there are multiple stationary source emissions above 1,000 tons per year, and 3) there are numerous ambient monitors measuring 5-minute and 1-hour SO_2 concentrations. The second draft may also evaluate exposures in the remainder of Missouri and also include areas of Pennsylvania, West Virginia, and other locations with large SO_2 emission sources. What are the views of the panel regarding the appropriateness of these proposed additional locations and on the approach used to select them?

The initial case study location is a reasonable choice. The committee had diverse opinions regarding inclusion of additional locations. Since the current results are derived from a single scenario based on a set of assumptions and local conditions, the relative value of varying the local conditions vs. assessing the sensitivity to other model input parameters and distributions should be evaluated. Additional sensitivity analyses in the current location could be an alternative and the additional effort in these analyses may be lower; however this should not be the only consideration. Towards selecting additional locations, an assessment of the national diversity of SO_2 concentrations as driven by multiple sources (i.e. in areas with overlapping plumes) may help with the location decision. Several committee members were interested in seeing an urban area in the northeast or an area near Midwest power plants included.

The text should reflect that the scenarios investigated are for a small subset of the US population near relatively large SO_2 sources and state that the numbers produced should be interpreted accordingly.

Charge Question 3: Do Panel members have comments on the appropriateness and/or relevance of the populations evaluated in the exposure assessment?

The populations were viewed to be generally appropriate and relevant. The most highly exposed populations will live near sources; it will be helpful to quantify the size of these populations since many sources are in rural areas. With respect to the table of age-specific asthma prevalence rates, the committee observed discontinuities and a few rates that were higher for females than males; these data should be checked. There is no provision for medication use in the modeling of the asthmatic population. Since the proportion of the asthmatic population that is routinely medicated could have implications for the health risk assessment, it may be worth addressing this point in the document.

Charge Question 4: To what extent are the approaches taken to model SO_2 emission sources technically sound and clearly communicated?

AERMOD is the appropriate concentration prediction tool if emissions-based modeling is determined to be the best option for obtaining the spatio-temporal concentration distribution. The model depends on a source inventory that may be incomplete and also omits many small and far away sources. This potential downward bias should be evaluated and an adjustment approach provided. Furthermore, there should be evaluation of whether upset conditions (high, short term emissions due to unusual events) from nearby sources are affecting predictions. Other alternatives to capturing the concentration distribution should also be considered.

The key consideration for assessing the quality of predictions for the purpose of this analysis is whether the predictions are comparable to real-world concentrations in level and variability. This analysis needs to capture the entire distribution (for the health risk assessment) as well as the peaks (i.e. the tails of the distribution, for the exposure risk assessment). Additional work is needed to assess the quality of the predictions.

Charge Question 5: Human exposures were modeled using APEX to simulate the movement of individuals through different microenvironments. Do panelists have comments on the microenvironments modeled?

APEX is an appropriate tool for the exposure model, and the approach represents current best practice. Generally the micro-environments and parameters chosen seem appropriate. Given the need to capture 5-minute peaks, the approaches to predicting 5-minute outdoor concentrations and to modeling infiltration are critical; these should be reassessed and in addition they should be given high priority for evaluation in sensitivity analyses. Description of the simulated exposure data should identify the times and locations where most of the peak exposures occur. If the higher exposures are occurring indoors, the assumptions for the air exchange rate and air conditioning prevalence may be particularly influential.

Given the continued reliance on APEX, EPA should ensure there is further evaluation and improvement of this model across a range of conditions and pollutants. Ultimately, as one component of this effort, human activity and representative population databases should be updated and expanded, including consideration of any unique features of activity patterns associated with asthmatics.

Health Risks:

Charge Question 1: What are the views of the Panel on the overall characterization of the health risk evidence for SO_2?

The draft REA draws on the ISA in selecting the outcomes and exposure-response relationships to be used. The principal reliance on the clinical studies of persons with asthma is appropriate for the quantitative analysis. There is a clear documentation of a causal association, and the exposure-response relationship has been characterized with reasonable certainty. The Panel also recommends that the Agency give consideration to the size and susceptibility of the subpopulations of persons with asthma susceptible to SO_2 at various concentrations. This sensitivity analysis should be carried out in a framework based around the context set by the epidemiological studies.

Charge Question 2: The characterization of health risks focuses on potential health benchmark values identified from the experimental SO_2 human exposure literature on lung function with accompanying respiratory symptoms. What are the views of the Panel on using potential health benchmarks from this literature to characterize health risks?

The Panel concurs with the use of the clinical studies to derive benchmarks for characterizing the health risks of SO_2 exposure. In using these studies, the Agency needs to acknowledge the highly selective nature of the volunteers included in these studies, who may not adequately represent the full range of sensitivity present in susceptible populations. Additionally, these studies addressed SO_2 alone and hence do not replicate the general circumstances of exposure to ambient SO_2, which is a component of a complex mixture having particulate and gaseous components that may influence dose.

Charge Question 3: Do Panel members have comments on the range of potential health effects bench mark values chosen to characterize risks associated with 5-minute SO_2 exposures?

The Panel strongly concurs that the range of values to be considered should be extended lower than the proposed range of 0.4 – 0.6 ppm. The range of exposures emphasized in the ISA (0.4 - 0.6 ppm) has clearly carried over to the thinking used in the REA. Adoption of this range might leave substantial numbers of exercising mild asthmatics at considerable risk. Studies (such as those listed in Table 5 – 1 of the ISA) have shown that 5 – 20% of mild to moderate asthmatics experience moderate or greater decrements in lung function at SO_2 concentrations as low as 0.2 – 0.3 ppm. For ethical reasons severe asthmatics were not part of these clinical studies, but it is not unreasonable to presume that they would have responded to even a greater degree, although routine use of medication among this group could influence response. In addition, the epidemiological evidence shows emergency room visits and hospitalizations for respiratory illnesses associated with 24-hour SO_2 levels below the current standard (0.14 ppm averaged over at 24-hour period). Collectively, this evidence should lead to a conclusion that 0.2 ppm or even a lower level of short-term exposure is an appropriate lower bound value for EPA's benchmark analysis.

Charge Question 4: To what extent is the assessment, interpretation, and presentation of initial risk characterization results technically sound, clearly communicated, and appropriately characterized?

The assessment, interpretation, and presentation of initial risk characterization results are technically sound, clearly communicated, and reasonable for a first draft. Attention has been directed at potentially susceptible subgroups. Additionally the basis for selecting the counties for the substantive characterization for benchmark health risks for 5-minute peak SO_2 exposure should be clarified. In particular, consideration of the representativeness of the locations needs to be included along with development of the approach that will be used to extend the results of the risk characterization nationally. Are there enough urban sites included? In characterizing risks, the Agency should give consideration to the possibility that SO_2 not only has effects of clinical significance for individuals but also has population-level effects that may be relevant to public health. Additionally, the Panel recommends an exploration of the sensitivity of the risk characterization to the model used for the concentration-response relationship. A variety of forms for this relationship can be assumed and different statistical models might be used to estimate the form of this relationship. Several members elaborate these possibilities in their comments (e.g. see comments of Drs. Hattis and Sheppard).

The discussion of uncertainty and variability remains completely generic. At this point, while there is extensive discussion of these matters with regard to exposure, and a probabilistic approach is described for addressing uncertainty in health estimates, the overall approach in the risk characterization remains to be specified.

Charge Question 5: The epidemiology literature will be used to qualitatively characterize SO_2-related health risks for health outcomes such as respiratory symptoms and emergency department visits and hospital admissions for respiratory-related causes. However, staff has judged that it is not appropriate to use the available SO_2 epidemiological studies as the basis for a quantitative risk assessment in this review. Do panel members have comments on this judgment and/or on the rationale presented to support it?

The Panel members had a range of views on use of the epidemiological data, ranging from considering the epidemiological evidence as qualitatively indicative of clinical morbidity resulting from the physiological responses documented in the clinical studies, on the one hand, to proposing that the epidemiological evidence should be used to develop a quantitative risk model, on the other hand. The Panel acknowledges that time constraints and resources may limit the approaches that can be taken by EPA for risk estimation. Regardless, the Panel members were not certain as to the nature of the qualitative risk characterization that will be performed. The epidemiological studies cited in the ISA address SO_2 as a component of a complex mixture and under exposure circumstances in which it may be a surrogate for other components of the air pollution mixture.

The CASAC Sulfur Oxide Panel was pleased to review the first draft of the REA and provide advice early in the development of this important document. We look forward to receiving the Agency's response and reviewing the second draft on December 17–18, 2008.

Sincerely,

Dr. Rogene Henderson
Chair
Clean Air Scientific Advisory Committee

Enclosure A: Roster of CASAC Sulfur Oxides Primary NAAQS Review Panel
Enclosure B: Compilation of Individual Panel Member Comments on EPA's *Risk and Exposure Assessment to Support the Review of the SO_2 Primary National Ambient Air Quality Standards*, July 2008)

Enclosure 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

[*] Did not participate in the July 30-31, 2008 review of the first draft REA.

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, 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.

Enclosure B: Preliminary Individual Comments on the *Risk and Exposure Assessment to Support the Review of the SO2 Primary National Ambient Air Quality Standards* (First Draft) from the Clean Air Scientific Advisory Committee (CASAC) Sulfur Oxides Primary National Ambient Air Quality Standards (NAAQS) Review Panel

Mr. Ed Avol

Comments on SOx REA – 1st Draft

The compilation of current thinking regarding the performance of a risk and exposure assessment to support the review of the SO_2 Primary National Ambient Air Quality Standard (NAAQS) should be descriptive, informative, logical, and clear. The first draft document provided for review demonstrates most, but not all, of these attributes. There is a great deal of information described within the document (although some of it seemed like unnecessary duplication of the ISA to me), and the general format is somewhat logical. However, I had trouble at times with the written clarity of the document. Several sections became fully immersed in detailed discussions or documentation of procedural steps to operate a model or perform an analysis, rather than describing the general approach. In my judgment, the more complete operational details of the model or application should have been relegated to an appendix or archive, and not be a part of the main text.

In my opinion, the pages and pages of so much detail (in the way of operational aspects, such as the text provided in Chapter 6) tended to blur the overall flow and logic of the document.

Moreover, there still appears to be a provincial perspective of accepting US and Canadian epidemiologic studies for consideration, but relegating other international studies (regardless of pedigree) as "…supportive evidence…" If the studies have withstood critical peer review and are published in well-recognized and respected journals, why should they not be considered equally?

With regard to the exposure perspectives presented in the document and the air quality analytical decisions (locations, data, approaches, etc), one question might be whether we are looking forward or backward in thinking about public health and exposure potential. The selection process for station data usage and exposure orientation is almost wholly guided by proximity to or downwind trajectory from power generation plants. In general, this is probably appropriate, given the currently understood sources and source strengths. However, these are typically located in rural areas away from major metropolitan areas and populations. It might be insightful to also consider other sources -- for example, seaport operations, where bunker fuels containing tens of thousands of ppm sulfur are routinely emitted, or rail, where fuels can contain hundreds or thousands of ppm sulfur, or large cumulative concentrations of traffic emissions (on-road vehicle sources individually may be quite low in sulfur emissions, but collectively may be a substantial line or area source). These activities and exposures tend to be in more urban and heavily populated areas.

Characterization of Health Risks
(1) (comments on the overall characterization of health evidence for SO2)

I generally found the presentation and characterization of the health evidence to be reasonable and appropriate. However, I am a little concerned that the threshold for consideration of data for the risk assessment may be unduly high or restrictive. While I agree with the general conclusion that the data regarding SO2 exposure for endpoints other than respiratory morbidity are not sufficient to infer a causal relationship, one wonders if the combined weight of multiple "not quites" or "almosts" should somehow count for something in the aggregate risk assessment. The controlled-exposure chamber data clearly support consideration of observable and clinically significant health outcomes at or below 0.2ppm SO2, and the epi data, in toto, paint a picture of clinically relevant detrimental health outcomes in a appreciable portion of the sensitive population.

(2) (use of clinical exposure data on SO2 to characterize health risks)
The use of clinical exposure data on SO2 to characterize health risks seems appropriate. Admittedly, the population studied in clinical research is small, somewhat self-selected, and generally biased towards increased interest/motivation in health and reduced severity of existing disease. Regardless, the observational data available from these numerous investigations are invaluable in establishing the potential for actual manifestation of specific health outcomes. The observed effects in these respiratory-compromised individuals legitimately raises the potential that other more individuals in the general population, with more severely compromised respiratory function who would not or could not participate in controlled chamber studies, are likely to be at risk to low-level (less than or equal to 0.2 ppm) and short-term (minutes to an hour) exposures to SO2.

(3) (comments on the range of potential health effects benchmark values chosen to characterize risks associated with 5-min SO2 exposures)
A number of controlled-exposure (clinical chamber) studies from the 1980s (primarily from US researchers at UCSF, Rancho Los Amigos Medical Center in Los Angeles, and the USEPA in Chapel Hill) demonstrated and confirmed the almost-immediate bronchoconstrictive effects of inhaled SO2 at levels in the 0.4-0.6ppm range, and also documents some responders down to and below 0.2 ppm. Subsequent clinical studies in the ensuing decades, though fewer in number and scope, have generally re-confirmed or extended these findings. Thus, the underlying evidence for a proposed range of health effects benchmark values seems available, corroboratory, and sufficient to support the consideration of a range of values for risks, in the 0.2 ppm vicinity, associated with 5-min SO2 exposures.

The more difficult issue to assess is the health implication of such short-term exposures, since many of the observed effects seemed to have declined, partially reversed, or resolved within a half hour or so of initial clinical exposure, even in the face of continuing exposures. Epidemiologic information regarding a range of health outcomes (including asthma-related symptoms, ED visits, and hospitalizations) as the result of short-term (5min to 24hr) exposures are less

consistent and convincing, but in the aggregate, do suggest an excess toll on human respiratory health. Therefore, it does appear that based on the available evidence, there is a susceptible and vulnerable population of people at risk from short-term exposure to SO2.

(4) (judgment on assessment, interpretation, and presentation of risk characterization results)
The assessment and presentation of the risk characterization results, based on ambient air quality and various permutations of peaks, peak-to-means, and other indices of exposure, seems extensive. In all of the presented detail, however, the clarity and summary points of interpretation are lost or rarely made. The Chapter 6 presentation is detailed and extensive, with page after page of plots and tables, but what is missing is a clear and succinct summary of what has been established by virtue of all the presentation by the end of the chapter.

Since the evolution of this format of ISA and REA is still in its infancy, it might be worth considering a slightly different format for presentation of the individual sections. The summary conclusion of each section could be stated in underline or bold format at the outset of each chapter section, and then the supporting material for the stated claim could be provided. This would have the advantage of clearly showing and stating the point of the ensuing presentation, discussion, or data. Alternatively, there needs to be additional effort made in the document to clearly state the summary conclusions in an accessible manner for document users.

(5) (comments on staff determination that SO2 epi data is not appropriate for quantitative risk assessment)
The staff recommendation that the available epi data is not sufficient to make a quantitative risk assessment is based, in part, on the determination that "…staff recommends primarily relying on US studies." (line 21, p167). The basis for this decision (to primarily use US studies only) is one that merits additional consideration, scrutiny, and potential reversal. Well-designed and executed studies are not limited to (or necessarily a boundary condition of) US-based studies. Unadjusted confounding variables and confounding exposures, lack of complete and precise study details, and well-constructed and appropriately performed statistical analyses challenge both American and foreign researchers. More useful activities would be to (a) identify specific gaps in available information needed for critical public health decisions, and (b) move aggressively to provide the necessary funding to obtain that information.

A separate question to be addressed is the relative level of staff comfort with regard to weight of evidence providable by epi data per se, compared to more controllable (but more artificial) exposure scenarios such as those utilized in clinical chamber work or animal toxicology. The trade-offs between real-world exposures of unrestrained mobile populations and lack of control or more complete understanding of those exposures have been noted and discussed on several occasions, but how to effectively and appropriately exploit the full value

of community or population-based studies to assess and protect public health is a critically important issue that should be explicitly resolved by staff so that the most appropriate judgments can be reached using the widest possible range of available, credible, and relevant data.

Dr. John Balmes

Comments on SOx REA – 1st Draft

2. The characterization of health risks focuses on potential health benchmark values identified from the experimental SO_2 human exposure literature on lung function with accompanying respiratory symptoms. What are the views of the Panel on using potential health benchmarks from this literature to characterize health risks?

In contrast to my opinion regarding the NOx Risk and Exposure Assessment, I support the staff decision to use the experimental SO_2 human exposure literature on lung response and respiratory symptom responses in subjects with asthma. This literature is sufficiently extensive to provide the basis for a quantitative risk assessment. I concur with staff's judgment that while the epidemiological literature shows relatively consistent associations with asthma outcomes (respiratory symptoms in children, emergency department (ED) visits and hospitalizations in children and adults), this literature is not sufficiently robust to support a quantitative risk assessment. That said, I endorse the staff's plan to use the data from recent U.S. and Canadian epidemiological studies of SO_2 and ED visits/hospitalizations to "qualitatively assess the range of SO_2 air quality levels that are associated with these endpoints."

Staff has reviewed the relevant controlled human exposure studies and selected health benchmark exposure values from those studies that reflect the potential for adverse effects in most asthmatic patients. Symptomatic bronchoconstriction will occur in a substantial proportion of such individuals when exposed for 5-10 minutes to concentrations of SO_2 between 0.4-0.6 ppm during exercise. However, it is also evident in the data from several of the controlled human exposure studies that some individuals responded with symptomatic bronchoconstriction from short-term exposure to 0.2 ppm. Thus, I recommend that 0.2 ppm also be considered as a health benchmark exposure value. As noted in the draft REA document such effects of SO_2 in controlled human exposure studies are coherent with the associations between ambient SO_2 and asthma outcomes reported in the epidemiological literature.

<u>Specific Comments</u>

p. 14, line15 should be "their" instead of "there".

p. 16, lines 14-17 The study by Winterton et al. to which this sentence refers found an association between the homozygous wild-type allele for a common polymorphism in the promoter region of TNFα (-308 G/A). The homozygous wild-type would be AA. This sentence should specify the specific polymorphism studied because there are other polymorphisms for TNFα.

p. 27, line 11 should be "…for boys *in a* Toronto, ON *study (*mean 24-h…"

p. 27, line 12 should be "to these *hospitalization* studies…"

p. 71, line 3 should be "…less *than* 1%..."

p. 135, Table 7-7 title should be "…children *in* the Midwestern U.S."

p. 136, Table 7-8 title should be "…adults *in* Missouri"

p. 147, line 4 should be "…dispersion modeled *concentrations* were…"

pp. 155-157 There is no discussion in this Uncertainty Analysis section of the uncertainties related to using National Health Interview Survey (NHIS) data for the prevalence of asthma in children of different ages or Missouri Department of Health data for the prevalence of asthma in adults from different regions of the state. For example, NHIS data are representative of the country as a whole, but do not have sufficient geographic resolution to be used at the state level. That is why Table 7-7 gives prevalence data for the Midwestern U.S. rather than Missouri.

p. 169, line 2 should be "…or *retrieve* …

Dr. Douglas Crawford-Brown

Comments on SOx REA – 1st Draft

These comments focus on Chapters 5 through 8 of the Risk and Exposure Assessment Draft, referring to earlier chapters only as they are needed.

I compared the conclusions in the early chapters to those in the ISA. The authors have been faithful to the primary conclusions from that earlier document. The same health effects, exposure durations (short-term) and sensitive subpopulations (asthmatics) are considered. It also places the same strengths and limitations, and hence sources of uncertainty, on personal exposure estimates found in the ISA.

In previous reviews of NAAQS assessments, including the recent one on NOx which uses similar methodologies, I have approved the proportional roll-up or roll-down methods. I support, therefore, the use of this method in the current document. The authors should state, however, any assumptions implicit in this approach, such as whether regulated sources and non-regulated sources are equally affected by any change in the NAAQS.

As in the draft of the NOx REA, I agree that the adjustment of the benchmarks produces the same result mathematically as adjusting the air concentrations. But it makes no sense scientifically, and is likely to be attacked as such. The savings in processing time don't appear to me sufficient to justify a method that people will fail to understand as mathematically equivalent, and will make it appear that the EPA staff is willing to make calculations based on an assumption that effects occur at levels below the benchmarks.

I support what is essentially a hazard quotient in the assessment (although the term is not used, the procedure is identical to one using an HQ calculation). The one issue I would raise here is that the hazard quotient approach usually has a margin of safety built in through uncertainty factors, and the current assessment does not appear to have this margin built in. The next draft should at least make mention of this issue.

The authors have done a much better job than in the first NOx draft of describing the relationship between the three approaches examined in the report. They seem to have learned from the NOx reviews.

I support the use of APEX and CHAD for the purpose of performing the stochastic calculations for the Chapter 7 analysis. These models contain assumptions that are routine in EPA assessments and have found application in a wide range of settings. They have been fully vetted for the kinds of assessments performed here. There remains, however, the problematic relationship between ambient levels as measured at monitors and ambient levels at or near the points of exposure for populations. I realize there is not much that can be done about that

issue, because the monitors are located where they are and can't be changed for the purposes of this assessment. But I would like to see a slightly better description of the implications of this problem for overall uncertainty.

As my expertise does not extend to air quality modeling, I can't comment on the adequacy of AERMOD for these purposes. It is a modeling package that has been used extensively in past EPA assessments, including the NOx assessment, and so I will assume here that it has been vetted. But I leave further vetting to other members of CASAC.

Assuming the air modelling can be performed adequately (and again, I will leave it to other CASAC members to comment on this in a more informed way), then the subsequent steps in Chapters 7 and 8 are reasonable. The development of the longitudinal activity sequences is a sophisticated piece of work, being state-of-the-science. The stochastic sampling methodology is reasonable and employed commonly at the EPA. The assumptions going into the sampling are adequately described. The microenvironments are both the correct ones to model given current data and well executed in the assessment steps (with a caveat about whether they correctly model the activities of asthmatics, which I note later in this review).

I found the characterization of results throughout informative and simple to follow. They walk the reader through the relevant findings. The one thing that continues to concern me is that I don't know how the results are to be used in any policy decision. For example, how many individuals, with how many exceedences, would count as acceptable or unacceptable in any decisions? I suppose it will be argued that those are policy concerns, not scientific ones, and that the only job of the REA to present these numbers. But I still expected to see at least some mention of this issue rather than leaving it entirely in the hands of policy staff and administrators.

I found it difficult to follow the uncertainty analyses, or at least to understand the magnitude and implications of any one source of uncertainty. I expected to see some statements, even if qualitative, about the uncertainty in the various risk results (e.g. uncertainty in number of people above a benchmark, percent of asthmatics experiencing a high exposure day one or more times). This aspect can be greatly improved.

I end with a comment I have made in other settings of CASAC, including in my review of similar methodologies for the NOx REA. The modelling performed here starting with Chapter 7 is impressive and represents state-of-the-science. But I worry that it may be hide a false sense of confidence in these results, which I take to be quite uncertain. There are many, many assumptions built into the assessment. At the moment, I think of the results as a kind of scenario analysis, and not necessarily an accurate reflection of actual exposures and risks in the US population. The methods may be pushing the current analytic ability too far.

It is for this reason that I believe the results of Chapter 9 will be quite important once they are produced. I realize the problems with epidemiology studies, but it seems to me there are equal uncertainties in the exposure assessments in Chapters 7 and 8. I think of the relationship between the epidemiology and clinical studies as one between Exact Questions, Approximate Answers, and Approximate Questions, Exact Answers. By this, I mean that the clinical studies ask a question (how do people respond when in a clinical setting?) that only approximates the one we want to ask (how do people respond in the natural setting?), but give a rather precise answer to that approximate question. Epidemiological studies address exactly the question we want, but provide only an approximate answer. I am not sure which approximation I prefer. In the end, perhaps the current results of this REA and those of the Chapter 9 assessment will need to be used as bounding answers. We will need to discuss this in more detail at the CASAC meeting.

Some Specific Comments:

Page 10- I don't see how the section Scenarios for the Current Assessment actually specifies scenarios. I was looking form greater detail here.

Chapters 2, 3 and 4 need headings, or at least introductory paragraphs, stating that these are reviews of the ISA conclusions. The Introduction says they are, but the reader may not remember that when reading the subsequent chapters..

The conclusions of Chapters 2, 3 and 4 are consistent with the draft ISA. Short term exposures and morbidity is the only association judged sufficiently strong in both documents.

0.4-0.6 ppm is identified in clinical trials to result in a substantial fraction of exercising asthmatics to have significant decrements in lung function, for 5-10 minute peaks. Why not just set the standard in that range, then? What is the purpose of the rest of the assessment? Is it only to explore the answers within that range? And if so, do the answers developed really allow us to differentiate the acceptability of a 0.4 ppm standard from one at 0.6?

I didn't review the part of Chapter 6 associated with air quality monitoring. I agree that application of PMR values is OK, but I can't comment on the empirical validity of these. I also am not convinced that the variability distributions used are valid in the tails of the distributions, which I suspect will affect the results.

On Page 71, I am not sure what is intended by the analysis of the impact of reducing the number of monitors. Why was this assessment done? I'm sure there is a reason, and suspect my inability to see it is related to my lack of understanding of this area, but some explanation would be good.

The uncertainty analysis in Chapter 6 has at least identified the major issues of uncertainty. Given that it is so qualitative, and doesn't involve any formal uncertainty analyses, it is hard to understand what the reader is to take away from this. Why were the results not run several times with at least some changes in parameter inputs, to at least get a sense of sensitivity? Still, there is probably no way to do a regular quantitative uncertainty analysis given the complexity of the calculations.

Chapter 7: APEX is the correct model for exposure. Not convinced it can model asthmatics well, however, so the assumption seems to be that they behave as the rest of the population in the CHAD database. Am I correct that this is assumed, and what are the implications on uncertainty if this assumed? I suspect asthmatics are less likely to go outdoors and play, especially during bad air quality days.

I applaud the use of decision trees on page 121. This kind of tree helps the reader understand the process used here. There are many places in the document where a similar tree would have been useful.

Is the assumption that one individual with N exceedences is the same (in terms of degree of adversity) as N individuals with 1 exceedence?

I'm not sure what to make of Tables such as 7-14. What would constitute large or small numbers? What is the criterion for this judgment? Or is the intent just to provide the numbers and let someone else decide in the policy branch? And how do we interpret a table which is both number of people and the number of exceedances per individual? There just seems to me to be too much flexibility in interpreting these tables.

The sensitivity analyses in Chapter 7 are better than in Chapter 6, although again this is not a full uncertainty analysis.

In Chapter 8, I am generally supportive of the approach. However, in the end, one will still be left with looking at the number of people above a given decrement of lung function or other metric, and so the logic will be the same as in Chapters 6 or 7. The only difference is that a 0.4-0.6 threshold will be replaced with a threshold based on level of decrement. I don't see this adds anything.

Dr. Terry Gordon

Comments on SOx REA – 1st Draft

Characterization of Health Risks (Chapters 3, 4, 5, 6, 7, 8, 9):

1. What are the views of the Panel on the overall characterization of the health evidence for SO2? Is this presentation clear and appropriately balanced?

The characterization of the health evidence was presented in a clear and balanced approach. The document is improved in style and clarity from the previous development document and is better in many respects than the first draft of the NOx REA.

2. The characterization of health risks focuses on potential health benchmark values identified from the experimental SO2 human exposure literature on lung function with accompanying respiratory symptoms. What are the views of the Panel on using potential health benchmarks from this literature to characterize health risks?

The choice of these benchmarks is appropriate and the uncertainty factors surrounding this data base were appropriately described with one possible exception. It must be noted that for health and ethical reasons, the clinical studies which form the basis of this assessment did not utilize moderate to severe asthmatics in the 5-10 minute exposure protocols. Therefore, the severity of pulmonary function decrements and asthmatic symptoms may be underestimated for the more severe asthma phenotype. EPA should present information regarding the relative numbers of mild, moderate, and severe asthmatics that make up the population of the U.S. and consider how these potentially more susceptible severe asthmatics may be affected by short term ambient exposure to SO2. Admittedly, the majority of the clinical studies were conducted in the mid-1980's. The subject criteria, medications, and disease severity classifications have changed since that time and, therefore, the uncertainty discussion on how well these subjects represent today's asthmatic population in the U.S. could be expanded.

3. Do panel members have comments on the range of potential health effects benchmark values chosen to characterize risks associated with 5-minute SO2 exposures?

The range of benchmark values in the exceedance calculations for exposed asthmatics utilized 0.4 ppm as the cut-off for health effects and, although this is out of my area of expertise, it was unclear why this was done in light of the health risk assessment which, utilizing probabilistic math, goes down to 0.2 ppm.

4. To what extent is the assessment, interpretation, and presentation of initial risk characterization results technically sound, clearly communicated, and

appropriately characterized?

The risk characterization is quite clear and technically sound.

5. The epidemiology literature will be used to qualitatively characterize SO2-related health risks for health outcomes such as respiratory symptoms and emergency department visits and hospital admissions for respiratory-related causes. However, staff has judged that it is not appropriate to use the available SO2 epidemiological studies as the basis for a quantitative risk assessment in this review. Do panel members have comments on this judgment and/or on the rationale presented to support it?

Although the epidemiology studies may not lend themselves to easy quantitative risk assessment, they are quite important despite the potential confounding by co-pollutants. In light of the positive findings in children and older adults, staff should make every effort to seriously consider these epidemiology data in a quantitative assessment, particularly if the qualitative assessment warrants such a step.

Minor Comments:
Page 14, line 15 – substitute 'their' for 'there'
Page 16, line 16 – There are many different alleles/polymorphisms for TNF so it is unclear which 'wild-type allele' is being referred to here (I believe it's the -308 polymorphism).
Page 71, line 3 – change 'thank' to 'than'
Page 161, line 10 – Should 'for' be added after 'model'?
Page 161, lines 20-22 – The refractory period has not been shown to last for a significant amount of time. Because Sheppard et al (1983) only looked at sulfur dioxide tolerance up to 90 minutes, a repeat exposure at 10 hours after the first 0.4 ppm exposure, for example, could cause a response. Although not identical challenges, it has been shown that tolerance to exercise-induced asthma is lost by 4 hours after the primary exercise challenge (Edmunds, 1987).

Dr. Dale Hattis

Comments on SOx REA - 1st Draft

Air Quality Information and Analyses (Chapter 6):

1. We have evaluated SO_2 air quality throughout the United States, using all available 5-minute and 1-hour ambient monitoring data for years 1997 through 2007. To what extent are the air quality characterizations and analyses technically sound, clearly communicated, appropriately characterized, and relevant to the review of the primary SO_2 NAAQS?

The basic approach of doing a set of empirical distributions of peak to median ratios based on a large database stratified by coefficient of variation (3 strata) and average SO2 level (5 strata) is reasonable. The only quibble is whether the number of strata selected for the two variables is the best choice. This could be tested by running a parallel analysis or two with greater numbers of strata of each type and comparing the bias and variability of the predictions vs the observations of peak levels with the base case analysis provided in the current document.

My major problem with Chapter 6 is its exclusive focus on quantifying exceedances of the very high health benchmark values (400 ppb and above). As I illustrated in my comment on the plan for the REA in the previous CASAC SO2 meeting, the problem of SO2 asthma responses is not well summarized by looking at a few localized sites where there are simultaneously very high concentrations (from local sources) and members of a sensitive subgroup known to react to those concentrations by direct clinical observation. In fact the problem needs to be analyzed as a combination of geographic/temporal variability in exposure levels combined with interindividual variability in sensitivity. In fact, based on the lognormal distribution of 1-hour ambient SO2 levels and the distribution of individual sensitivity thresholds observed by Horstman et al. (1986) my earlier calculations indicated that only about 22% of the total events causing asthmatics to endure a 100% increase in specific airway resistance would occur at concentrations of 400 ppm and above. Therefore it would be more reasonable for any subsequent version of the REA to include exceedances of at least a few lower SO2 levels. I have recently updated this analysis to factor in the smaller ED50 and slightly greater interindividual variability indicated by the Linn et al. papers included in the updated ISA. This revised analysis indicates that only 11% of the overall population asthma-exacerbation effect could be expected to occur at over 400 ppb. About 50% of the expected effect is likely to occur at concentrations of 160 ppb and below.

As a further step in this analysis I have fit lognormal distributions to the exposure levels derived in the new REA. It can be seen in Figures 1 and 2 that lognormal distributions do not fit perfectly to these results—if anything the lognormal fits tend to underestimate the frequency of very high exposure levels. Despite this,

using the lognormal distributions of 5 minute maximum exposures to exercising asthmatics for current emission sources in Missouri, less than 2% of the overall asthma exacerbation effect is expected to occur below 400 ppb. About 50% of the expected to occur at exposure levels of 120 ppb and below. The further diminished importance of vary high exposure levels results from a lower overall variability and higher geometric mean exposure in these 5-minute exposure estimates relative to the previous estimates for the national distributions of 1 hour concentration levels at ambient monitors.

Figure 1

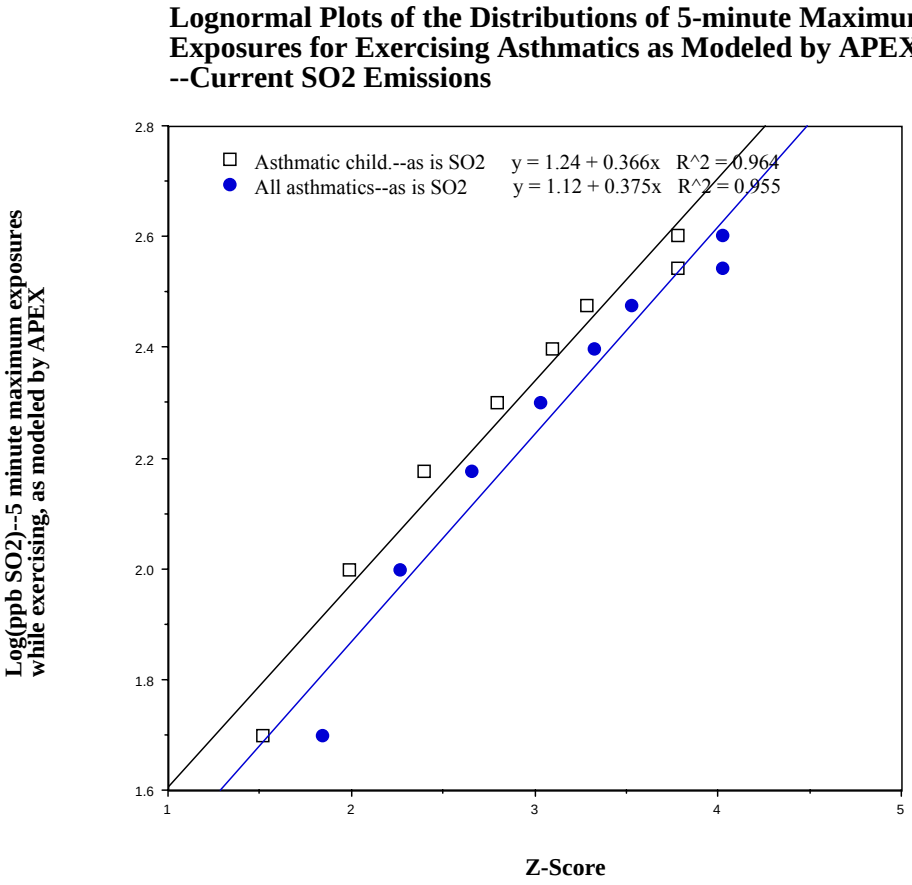
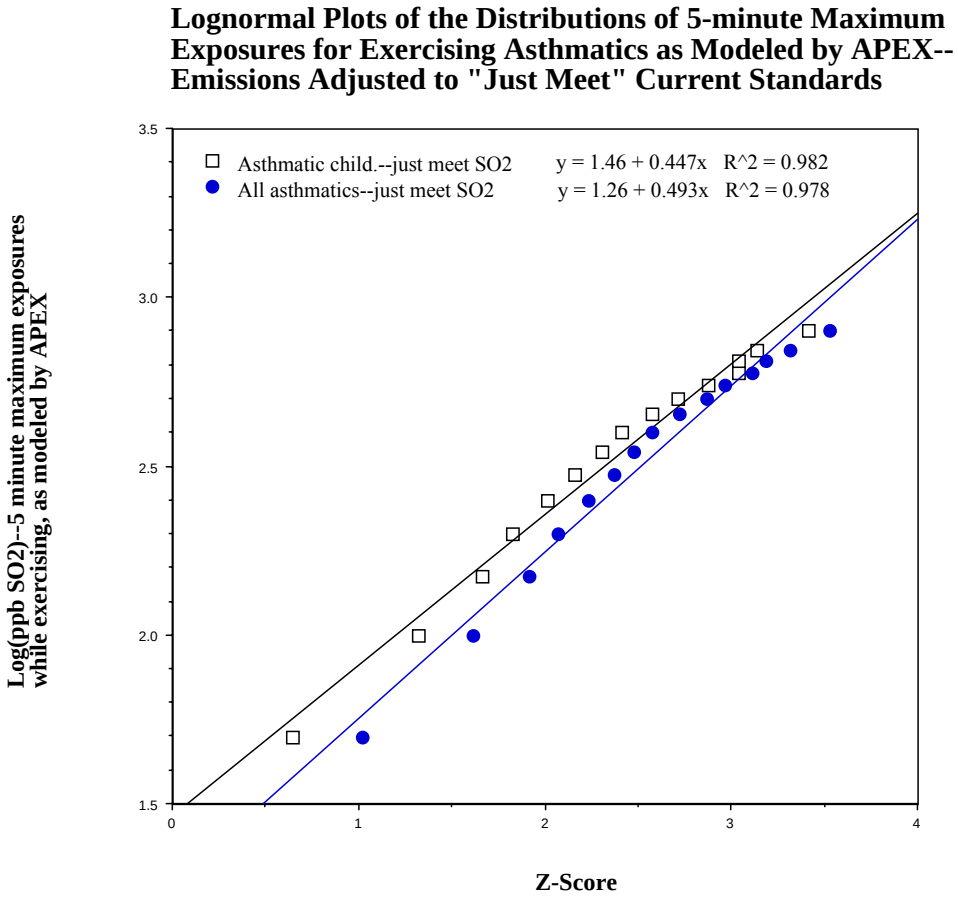


Figure 2



2. To what extent are the properties of ambient SO_2 appropriately characterized, including ambient levels, spatial and temporal patterns, relationships between various averaging times, and the relationship between ambient SO_2 and human exposure?

3. Twenty locations were selected for detailed analyses, using ambient SO_2 monitoring data for years 2002-2006. What are the views of the panel regarding the appropriateness of these locations, the time period of analysis, and the approach used to select them?

These seem reasonable to me.

4. In order to simulate just meeting either the current 24-hour or annual standards, staff adjusted SO_2 air quality levels for the years 2002-2006 upwards in all but one location. Ambient monitoring data in North Hampton County PA were above

the 24-hour standard in the year 2006 and were therefore adjusted downward. To what extent is the approach taken technically sound, clearly communicated, and appropriately characterized?

These seem reasonable to me.

5. What are the views of the Panel regarding the adequacy of the assessment of uncertainty and variability?

The document appears to do a reasonable job at this.

Exposure Analysis (Chapters 2, 7):

1. To what extent is the assessment, interpretation, and presentation of the initial results of the exposure analysis technically sound, clearly communicated, and appropriately characterized?

I have a number of problems with the analysis and its summarization. Specifically:

p. 142—the method for assessing indoor exposures assumes there is only one peak outdoor level per hour of exposure—all the rest of the 5 minute periods will have an average level assigned, calculated after excluding the peak 5 minutes. This will mean that indoor exposures will have much lower peak levels as slow air exchange rates will effectively dilute the 5 minute peaks toward the hourly averages.

p. 145—removal rate distributions are represented as uniform and rather high—"Resulting estimates were as follows; morning: 4.9 – 19.8 h^{-1} and afternoon: 3.4 – 9.8 h^{-1} " How are these derived from the data? Reproduce some summary of the data and analysis from the cited paper of Grontoft and Raychaudhuri, 2004.

Grontoft T and MR Raychaudhuri. 2004. Compilation of Tables of Surface Deposition Velocities for O3, NO2 and SO2 to a Range of Indoor Surfaces. Atmos Environ. 38:533-544. 27

In general I disapprove of the use of uniform distributions because they imply zero probability of occurrence of values outside the designated range. There is usually no good reason to assume this. The data underlying these distributions must be fully described, in addition to reasonable methodology to derive distributional inputs for this factor for use in APEX. Because the indicated ranges will lead to very large reductions in expected SO2 indoor concentrations, this is a key issue for the modeling of indoor exposures.

p. 157—the uncertainty section does not discuss the uncertainties in the indoor removal rate—likely a very influential variable, at least for indoor exposures

2. The draft risk and exposure assessment evaluates exposures in selected locations encompassing a variety of SO_2 emission source types in the state of Missouri; these areas were chosen as an initial case study since 1) air quality measurements indicated numerous exceedances of 5-minute benchmark values, 2) there are multiple stationary source emissions above 1,000 tons per year, and 3) there are numerous ambient monitors measuring 5-minute and 1-hour SO_2 concentrations. The second draft may also evaluate exposures in the remainder of Missouri and also include areas of Pennsylvania, West Virginia, and other locations with large SO_2 emission sources. What are the views of the panel regarding the appropriateness of these proposed additional locations and on the approach used to select them?

These seem reasonable to me.

3. Do Panel members have comments on the appropriateness and/or relevance of the populations evaluated in the exposure assessment?

These seem reasonable to me.

4. To what extent are the approaches taken to model SO_2 emission sources technically sound and clearly communicated?

P. 125 discusses the approach for using Aeromod for deriving outdoor concentrations for input into the exposure model as follows:

"As discussed above, as a first approximation point sources at major facilities were assumed to represent the SO_2 emissions throughout Missouri20, where major facilities were defined as those with SO_2 emissions totals exceeding 1,000 tpy. Nationwide, there are 918 major facilities and 10,651 associated stacks, according to the 2002 NEI. Within Missouri, 281 major facility stacks were identified, but only 115 of these stacks have greater than or equal to 1.0 tpy SO_2 emissions in the 2002 NEI. Each of these stacks was paired to a surface meteorological station, defining its modeling domain. These are the final list of stacks identified in Table 7-1, above. **"**

It seems to me this guarantees an underestimation of emissions as the concentrations resulting from many smaller sources within and outside 20 km of the major sources will be omitted. The document as it stands does not seem to provide an approach for adjusting the estimated ambient outdoor concentrations upward to reflect this source of systematic bias.

Some comparison between predicted and measured concentration distributions for a few monitors n Green County, Mo. is provided in Figure 7-3. The figure does not provide the detail needed for quantitative comparison that a tabular comparison would; and the discussion is vague and qualitative—saying mainly that the distributions seen at the monitors are "bounded" by the modeled values. The EPA should develop a procedure to quantitatively adjust the modeled distributions to distributions observed at some reasonably representative set of monitors, as was previously suggested for the NO2 analysis.

5. Human exposures were modeled using APEX to simulate the movement of individuals through different microenvironments. Do Panel members have comments on the microenvironments modeled?

Characterization of Health Risks (Chapters 3, 4, 5, 6, 7, 8, 9):
1. What are the views of the Panel on the overall characterization of the health evidence for SO_2? Is this presentation clear and appropriately balanced?

2. The characterization of health risks focuses on potential health benchmark values identified from the experimental SO_2 human exposure literature on lung function with accompanying respiratory symptoms. What are the views of the Panel on using potential health benchmarks from this literature to characterize health risks?

As mentioned in my response above to question #1 of the air quality analysis, the distribution of individual sensitivities among asthmatics mean that there are likely to be appreciable numbers of asthmatics who respond to 5 minute exposure to concentrations less than the lowest 400 ppb benchmark analyzed. At the very least a series of lower benchmark values should be included in parallel characterizations.

3. Do panel members have comments on the range of potential health effects benchmark values chosen to characterize risks associated with 5-minute SO_2 exposures?

As mentioned in my response above to question #1 of the air quality analysis, the distribution of individual sensitivities among asthmatics mean that there are likely to be appreciable numbers of asthmatics who respond to 5 minute exposure to concentrations less than the lowest 400 ppb benchmark analyzed. At the very least a series of lower benchmark values should be included in parallel characterizations.

4. To what extent is the assessment, interpretation, and presentation of initial risk characterization results technically sound, clearly communicated, and appropriately characterized?

5. The epidemiology literature will be used to qualitatively characterize SO_2-related health risks for health outcomes such as respiratory symptoms and emergency department visits and hospital admissions for respiratory-related causes. However, staff has judged that it is not appropriate to use the available SO_2 epidemiological studies as the basis for a quantitative risk assessment in this review. Do panel members have comments on this judgment and/or on the rationale presented to support it?

In a late presentation that did not make it in to the written REA we are reviewing, it now appears that the REA authors will derive a concentration-response function from the clinical data for asthmatics and therefore be in a position to quantify the extent of likely asthmatic responses to the full distribution of concentrations to which people are exposed under different SO2 NAAQS options. This is a giant step forward. However the current model under consideration is based on a 3-parameter logistic function with a firm upper limit of population response that is less than 100%. At least to provide sensitivity analysis for this model form, EPA should include calculations based on a probit model with one or two subpopulation modes. The probit model is based on an assumption that individual thresholds for response have a lognormal distribution in the population. The two-population model would hypothesize a mixture of two lognormals to describe the population distribution20of thresholds among asthmatics. I prefer lognormal distributions because they have at least some quasi-mechanistic justification: the likely possibility that there are many factors that contribute to differences among people in their individual thresholds, and the influences of these factors tend to act multiplicatively. The normal distribution of log(threshold) values then follows (approximately) from the central limit theorem. In my view it is better to choose model(s) with some mechanistic justification because projections to effects at lower concentrations than covered by the data depend crucially on the consistency of the model form chosen to represent the real causal relationship.

Dr. Ted Russell

Comments on SOx REA – 1st Draft

This document provides an analysis of air quality data and lays out the modeling approach EPA plans to use to calculate the number of individuals exposed to SO2 levels of concern in association with varying potential standards. The Draft begins with a history of the standard and overviews of SO2 exposure, at risk populations and health effects. The two main components of the current draft are the ambient air quality characterization for 5-minute exposures and the lay out of the exposure analysis. The Health Risk Assessment and Risk Characterization chapters are not yet fully developed. The current draft shows a significant amount of effort.

A starting comment is that the Introduction should lay out a road map for the document discussing what is being done and why. A second comment is that the while some of the document is relatively easy to read, other aspects are more difficult, and one is asking why are they doing this? How will they use this analysis? Was this necessary? The Overview of the Assessment (Section 1.2.1) is insufficient in this regard.

Chapter 2 on Human Exposure is quite brief, and is more properly titles an overview. Given the consideration of 5-minute levels in the risk characterization, it is curious that this concentrations at that averaging time are not even mentioned in Section 2.3. Further, is there really much concern about the instrument being used in regards to attainment demonstration? What is the reason for concern? The longest of the three paragraphs in Section 2.3 is on the PRB, which is not even used. This section should be more balanced and address the concerns addressed in the rest of the document.

Chapter 6, Ambient Air Quality and Benchmark Health Risk Characterization, is very dense at this point, at it is not even clear what is being gained from all of the analyses. As a first comment, this chapter needs to be cleaned up and streamlined, written with specific objectives in mind. Indeed, it appears that more analyses might have been presented/done than needed.

A primary objective of Chapter 6 is to provide an appropriate characterization of five-minute peak SO2 concentrations for use in exposure assessment. While some monitors do provide 5-minute average data, most only provide one-hour data. However, there are enough locations that provide both to develop relationships between 5-minute and one-hour average peak concentrations. The approach taken has many aspects of what I would deem appropriate, but could be improved, I think. In particular, the 5-minute and 1-hour average concentration data are derived from the same population of observations of SO2 concentrations at a single location, so a solid understanding of the distribution of pollutant concentrations, and correlations between 5-minute and 1-hour levels should

provide an avenue for deriving 5-minute peaks from 1-hour peaks. This is the approach they take in deriving a peak-to-mean ratio. Where I might differ in their analysis is that I would start from the assumption that the concentration data follow a log-normal distribution (this should be tested for individual sites as well as the population as a whole), and for each site, derive the geometric mean and standard deviation (GM and GSD). A COV uses the traditional standard deviation, which is based on the underlying population being normally distributed, which it is not. They could then analyze the relationship between the 5-minute and 1-hour GSD's. Assuming that they find as good of a relationship between the GSD's as they did between the COVs (and it would be difficult to think they would not, given the results for the COVs), they can then readily identify expected percentile values based upon the observed geometric means (care must be taken in how to treat below detection limit values). This would negate the need to do much of the Monte Carlo analyses they currently do. I would think that a very reasonable functional dependence of the 5-minute peak on the 1-hour maximum and 1-hour GSD can be found, and that this relationship can be used to estimate the maximum (or second, third, etc.) 5-minute peak level at each monitor. In essence, I am suggesting that they use that the concentration data likely are log-normally distributed to simplify and strengthen much of their current analyses. As part of this, they should develop the temporal correlation structure of the 5-minute data, as well as the correlation between 5-minute and 1-hour average data.

From Section 6.2.3.6 and on, the document gets dense, and it appears as though much of the analyses are not central to the objectives of the chapter. The motivation behind the analyses need to be better brought out with respect to how they will be used in the ensuing exposure and risk characterizations.

The APEX modeling chapter (Chapter 7) is much more readily understood than Chapter 6. The one issue that I am a bit uncomfortable with in this chapter, however, is the simulation of 5-minute SO2 peaks indoors. Given the inertia effect of the indoor air diluting the peak levels, the temporal correlation between outdoor 5-minute levels may be critical to correctly calculating the distribution of 5-minute average levels indoors. Currently, they assume that all of the other 5 minute periods ha d the same concentration. If one were to assume that there were more structure, e.g., that half of the concentrations were zero, and the other half at the peak, and further, that all of the peaks occurred together, one could get a higher peak level indoors (and that level would be very sensitive to the infiltration rate used).

Page 43 lines 27-28 "although though" should be corrected.

Page 45, line 17: "having an estimated" not "containing estimated".

Figure 6-7: Use a log scale for this type of graph.

Table 6-7: add "(ppb)" in the table

Page 68, line 14: This does not really seem to match what is in Fig. 6-16 (and the upper and lower rows of Fig. 6-16 are nearly the same, and no additional information is transmitted by the upper row).

Table 7-8: "… the Missouri" What???

Page 143, lines 12 and 13: Replace NO2 with SO2.

Table 7-12: Add units.

Section 7.9.2 and Tables 7-14 through 7-17: Add complete units, e.g., per year, etc.!

Page 71lines 2-3 "400 ppb at **any one** monitor was **between** 20 to 60 times a year … less **than** 1%..."

Figures 6-21 and 6-22: It looks as though there are fractional numbers here.

A first quibble is that the Introduction could be expanded to provide more of a picture of what was to come. A few paragraphs laying out the approach would be good, providing a flow of effort and information. Here they can define what models are to be used and why, as well as the specific outcomes of interest, and why. A second general comment is that the document is a bit uneven, with some sections being thorough and readily understood, while others lacked motivation and it was a bit difficult to see exactly what was done and why.

Page 86, line 12: The reference to the content of Fig. 6-28 is confusing.

Page 103, line 26 "…samplers **for short term averages**"

Page 103, Line 28 "…days, and 5-minute averages are never available."

Page 104, lines 3-7: Unclear what is being said (and why)

In response to the specific Charge Questions:

Air Quality Information and Analyses (Chapter 6):

1. We have evaluated SO_2 air quality throughout the United States, using all available 5-minute and 1-hour ambient monitoring data for years 1997 through 2007. To what extent are the air quality characterizations and analyses technically sound, clearly communicated, appropriately characterized, and relevant to the review of the primary SO_2 NAAQS?

As discussed above, I have a few concerns about Chapter 6. First, that Chapter does not clearly communicate what has been done, and why, and a large fraction of what is there ends up appearing to be of secondary relevance to this review. In many places, the figure and table captions need to be expanded to better convey what is being plotted/tabulated. Technically, the use of a traditional COV is questioned since it uses the traditional standard deviation, which is appropriate to characterize populations that are normally distributed. Primary air pollutant concentrations typically are log normally distributed, and thus one should log-transform the data first. That said, I can support the spirit of how they are finding 5-minute peak values given 1-hour data, just that I would look to start with using geometric means and standard deviations (taking care of how below detection limit data are treated). I realize it is late in the process, but this is where there is a mismatch between the ISA and REA in that the ISA has little on 5-minute average SO2 levels, but it is central to the REA analyses.

In addition to considering characterizing the distributions assuming they follow a log-normal distribution and developing the appropriate relationships and correlations between the 5-minute and 1-hour concentrations, I would look to streamline this chapter with the ultimate goal in mind: to characterize the distribution of peak SO_2 levels, particularly those above 400-600 ppb (at least for now). With that in mind, I would look to see what analyses are central to such. For example, consider Figure 6-8. Why is one concerned with having 3 different monitors in the county? What exactly is plotted (the figure caption is insufficient as to what each dot represents)? The discussion on page 49-50 does not help answer this question. I think the real question is independent of having three or more monitors. The discussion related to Table 4 is a bit opaque: how is the COV defined? How is the COV used? Finally, on page 54, one sees how multiple monitors are used (but it is still not apparent why this is a requirement), but this could have been rather simplified.

Section 6.4 starts off well, but then gets bogged down in analyses. (On the other hand, Table 6-5 should also have slopes from the regression, and I assume that in Figure 6-9, the RH Column is annual 1-hour average **Max** to be consistent with the left hand column). For example, the upper rows in Figures 6-10 and 6-16 provide little extra insight. Fig's 6-13 and 14 are informative.

In regards to Table 6-7, not surprisingly, the distribution of the modeled 5-minute maximums is not normal, so, again, a standard deviation is not an appropriate measure. Again, I would providing the geometric mean and standard deviations.

2. To what extent are the properties of ambient SO_2 appropriately characterized, including ambient levels, spatial and temporal patterns, relationships between various averaging times, and the relationship between ambient SO_2 and human exposure?

As noted above, most of the analysis uses statistics and characterizations for populations that are normally distributed, which they are not. There really is no solid analysis of the relationship between ambient SO2 and human exposure, except in the uncertainty section. Given the lack of analysis of this relationship, the uncertainty discussion seems out of place.

3. Twenty locations were selected for detailed analyses, using ambient SO_2 monitoring data for years 2002-2006. What are the views of the panel regarding the appropriateness of these locations, the time period of analysis, and the approach used to select them?

The twenty locations are fine. As noted above, it is not apparent that there was a need for having three or more monitors in a county was a necessary criteria.

4. In order to simulate just meeting either the current 24-hour or annual standards, staff adjusted SO_2 air quality levels for the years 2002-2006 upwards in all but one location. Ambient monitoring data in North Hampton County PA were abovethe 24-hour standard in the year 2006 and were therefore adjusted downward. To what extent is the approach taken technically sound, clearly communicated, andappropriately characterized?

The approach used was fine, though not overly well communicated. This latter paragraph (i.e., the charge question) actually brings clarity to what was done and why.

5. What are the views of the Panel regarding the adequacy of the assessment of uncertainty and variability?

At present, the analysis is qualitative. It should be a bit more quantitative in regards to the probabilities of having concentrations exceeding specified values, and the numbers of exceedences. It does reasonably well on the numbers, but could do a bit better on probabilities of a certain number of exceedences.

Exposure Analysis (Chapters 2, 7):

1. To what extent is the assessment, interpretation, and presentation of the initial results of the exposure analysis technically sound, clearly communicated, and appropriately characterized?

2. The draft risk and exposure assessment evaluates exposures in selected locations encompassing a variety of SO_2 emission source types in the state of Missouri; these areas were chosen as an initial case study since 1) air quality measurements indicated numerous exceedances of 5-minute benchmark values, 2) there are multiple stationary source emissions above 1,000 tons per year, and 3) there are numerous ambient monitors measuring 5-minute and 1-hour SO_2

concentrations. The second draft may also evaluate exposures in the remainder of Missouri and also include areas of Pennsylvania, West Virginia, and other locations with large SO_2 emission sources. What are the views of the panel regarding the appropriateness of these proposed additional locations and on the approach used to select them?

While these locations are appropriate, I (and I think the panel) will always be most interested in a national perspective.

3. Do Panel members have comments on the appropriateness and/or relevance of thepopulations evaluated in the exposure assessment?

They are fine to me.

4. To what extent are the approaches taken to model SO_2 emission sources technically sound and clearly communicated?

AERMOD is the appropriate tool if emissions-based modeling is decided to be the best route, though I am not sure that one needs to go that way. Might one rely on just the analysis of the observations? Does using AERMOD add an extra complication?

5. Human exposures were modeled using APEX to simulate the movement of individuals through different microenvironments. Do Panel members have comments on the microenvironments modeled?

While APEX is an appropriate tool to be used in this case, the continued reliance on APEX should push EPA to further evaluate the model across a range of conditions and pollutants. The lack of evaluation in this application is not comforting, though understandable given the limitations in measurements available. Also, given the task at hand, i.e., simulating 5-minute maximums, how infiltration is done is important. Also, it would be of interest to show where and when the exposures to > 400, 500 and 600 ppb occur. Do they happen in the home, at night, etc. This is an uncommon, short term, affect.

In Tables 7-14, 15: The number exposed above 0 should be all of the individuals, independent of number of exposures.

In regards to the uncertainty discussion, the treatment of the AER's and air conditioning prevalence could be very important if the 5-minute peak exposures above the thresholds are happening indoors. These sections may need to be bolstered if that is the case.

Dr. Donna Kenski

Comments on SOx REA – 1st Draft

General comments and responses to charge questions:
1 We have evaluated SO2 air quality throughout the United States, using all available 5-minute and 1-hour ambient monitoring data for years 1997 through 2007. To what extent are the air quality characterizations and analyses technically sound, clearly communicated, appropriately characterized, and relevant to the review of the primary SO2 NAAQS?

The 5-minute and 1-hour data has been exhaustively analyzed in Section 6, but it was not always easy to see the path being followed or the logic of the method pursued. I sometimes felt lost in the minutia, and had a hard time keeping all the pieces of this analysis in perspective. So it could use some additional clarification of the overall structure. Or maybe just some judicious editing with less detail an more summarizing— some suggestions for items that could be sent to an appendix are below. Other items were not explored as thoroughly as needed, however. Two of the first sections brought up issues that were never returned to; the duplicate dataset (6.2.1) and the distance from monitor to sources (6.2.2). For instance, where did the duplicate data enter into the QA process? I couldn't find it referred to again, after the first description, until Appendix A. It seems like this dataset should have been used to test the PMR model, but I couldn't see any indication of that. The analysis of the duplicates in the Annex was okay, but this particular part of the dataset could have been used more effectively in model validation. Likewise, the characterization of monitors by their distance from sources seems like information that could have been used to improve or inform the predictive model. The choice of COV as a predictive categorical variable is reasonable, but the REA could have benefited from a more comprehensive discussion of possible models and the rationale for that particular choice. Sec. 6.2.3.1 (Background) explains that peak concentrations are likely to be influenced by distance from sources and source characteristics, but the subsequent justification for the COV model was weak. Since data on source types, emissions, and distance from monitors were available, it is not clear why they were not explored at least briefly. More importantly, there is some discussion in Sec. 6.2.3.6 (Evaluation of Estimation Procedure) of model fit and some outliers. The poor model fit at 2 specific monitors is discussed as being perhaps a function of the proximity of the monitors to the nearby sources, or some unspecified characteristic of the sources that causes them to be poorly described by the statistical model. These two cases are then excluded to demonstrate improved agreement. But these two cases are among those that should have the closest scrutiny, since they are generating values at the extremes of the distribution. They should be examined in detail rather than discarded for the sake of showing better model performance.

2 To what extent are the properties of ambient SO2 appropriately characterized, including ambient levels, spatial and temporal patterns, relationships between various averaging times, and the relationship between ambient SO2 and human exposure?
Most of the comments above pertain to this question as well. In addition, the trend information in section 6.4.2 seems like it is of limited use in this analysis. Perhaps it belongs in the appendix? It is well documented that SO2 concentrations have been declining as a result of several regulatory programs. I'm not sure how those trends are helpful in interpreting the risk and exposure assessments that are made, or will be made, in this document. Some additional justification of this particular analysis would be helpful. Excluding the results for the Caribou ID monitor (p. 70) is another instance where an outlier is discarded that might be more useful if analyzed separately or in more detail to look at the reasons for its behavior. The exclusion of the Hawaii County data, on the other hand, is perfectly valid.

3. Twenty locations were selected for detailed analyses, using ambient SO2 monitoring data for years 2002-2006. What are the views of the panel regarding the appropriateness of these locations, the time period of analysis, and the approach used to select them?

The first 3 of the 20 locations are certainly good choices. It is not clear exactly why the remaining 17 were selected – i.e., why was it necessary to have 3 monitors in a county? Surely it is more important to have the highest-concentration monitors represented? I can't tell what impact this choice of monitors might have on the ultimate results of this analysis, but it seems as though it might be significant in terms of the number of potential exceedances. Consequently the selection rationale needs to be more completely justified, and/or some of the higher concentration monitors should replace the 3-monitor counties.

3 In order to simulate just meeting either the current 24-hour or annual standards, staff adjusted SO2 air quality levels for the years 2002-2006 upwards in all but one location. Ambient monitoring data in North Hampton County PA were above the 24-hour standard in the year 2006 and were therefore adjusted downward. To what extent is the approach taken technically sound, clearly communicated, and appropriately characterized?

I thought this approach was fine and clearly communicated.

4 What are the views of the Panel regarding the adequacy of the assessment of uncertainty and variability?

Frequently unclear. Section 6.2.1 starts off with a description of a dataset of duplicated measures that were used for quality assurance, but the rest of the document doesn't refer back to this particular dataset so it is hard to assess the level of QA with these data. Then later, in Sec. 6.2.3.6 (~ p. 45) it sounds like the

model estimates are compared with measured values at the same sites that were used to develop the model. Maybe I'm misreading this? Obviously an independent dataset should be used to evaluate the model. You could do this by reserving some fraction of the data for this purpose if the collocated duplicates are not suitable. In either case it is not clear whether the model has been evaluated with the appropriate set of data and that should be clarified. Overall, Sec. 6.5 did a nice job summarizing in a qualitative way the various sources of uncertainty. It would be nice to have a tabular, graphical or bullet summary of the various uncertainties described in section 6.5.

Specific comments:

p. 10, line 15 delete the 'is' after scheduled

p. 11 line 14 missing a period

p. 14 line 15 there -> their

p. 19 the bullet list of key conclusions is nice. In fact this whole introduction section was well done.

p. 20 line 5 as low as 0.4 ppm

p. 22 line 15 visits

p. 25 1st paragraph add the n for this study.

p. 25 line 28 really per 40 ppb? Or 10 ppb?

p. 28 lines 19-23 This sentence is unbearably long and should be broken up and/or reworded for clarity

p. 32 line 3 data were assembled

p. 35, figure 6-1 The labels in this figure don't seem right. How could hydro power contribute 30% of SO2 emissions? And in part B, electric power generation is allocated 2%, vs. fossil fuel power generation at 45%. Needs clarification.

p. 40, footnote reference to Fig. 4 should be Fig. 6-4

p. 42, line 17 1-hour measurements should be 5-minute?

p. 42, footnote Why was a uniform distribution used? It's not clear whether that's the appropriate choice here; provide some justification. Also, should be …was based on selection of a value…

p. 43, line 28 delete 'though'; change resultant to resulting

p. 50, line 8 delete 'both'

p. 57 Fig. 6-9 Add units to coefficient of variatiability

p. 59 Caption of this figure is a bit confusing, as it uses the words exceedance and benchmark interchangeably. Since exceedance has a specific regulatory implication, and that's not what's being discussed here, it would be better to stick with benchmark (this situation comes up in numerous places in the text and figures)

p. 64, line 27 series of figures

p. 65, Fig. 6-15 Add units to coefficient of variatiability

p. 68, line 4 re -> Fig.

pps 69 and 73: In both of these figures, the top row and bottom row are so similar that they can't be meaningfully distinguished from each other. Replot on a log scale to show the differences, if they are important (and here's another exceedance vs. benchmark confusion)

pps. 75-78, Figs. 6.20-6.23 It might be helpful to color-code the points by year and graphically make the point that the concentrations above the benchmark occurred only in early years.
Pps 84-86, Figs 6.24-6.26 This series of figures is not very effective as laid out. They would be better if the bars were side by side so we could see the change in magnitude between the as-is and adjusted values.
p. 100, line 18 …; however, it incorporates…
p. 100, line 29 delete data
p. 101, line 16 impact on
p. 102, line 1 delete those
p. 104, lines 1-3 fix incomplete sentence
p. 108-110, Figs 6.34-6.36 the symbols aren't really legible on these plots

Dr. Patrick Kinney

Comments on SOx REA – 1st Draft

Exposure Analysis Charge Question 5: Human exposures were modeled using APEX to simulate the movement of individuals through different microenvironments. Do Panel members have comments on the microenvironments modeled?

Overall, the approach taken by EPA in applying APEX to the SO2 exposure and risk assessment represents best available practice using currently-available modeling tools. The microenvironments chosen for inclusion, and the parameters assigned to each, are reasonable.

It is worth noting that the human activity data base upon which the modeling work depends represents a compilation of results from human activity surveys conducted between 1982 and 1998, and thus are 10 or more years old. EPA should consider updating these data periodically, both by summarizing results from more recent time/activity survey studies, and if necessary, by sponsoring new population-based surveys.

p. 114, line 23 through p. 115, line 2: this discussion is unclear.

p. 143, lines 12-14: this text is for NO2. Please edit for SO2.

Dr. Steven Kleeburger

Comments on SOx REA – 1st Draft

Characterization of Health Risks (Chapters 3, 4, 5, 6, 7, 8, 9):

3. Do panel members have comments on the range of potential health effects benchmark values chosen to characterize risks associated with 5-minute SO_2 exposures?

The potential health effects chosen for consideration that are consistent or in common between the ISA and REA documents include respiratory symptoms (e.g. wheeze, chest tightness, cough, substernal irritation), lung function (e.g. change in FEV1, sRaw, , decrements in lung function in the presence of respiratory symptoms, and cardiovascular parameters. Given the existing epidemiological, clinical, and animal model investigations of health effects related to 5-10 minute SO2 exposures, the selection of these health effects was reasonable. Moreover, the potential "affected individual" or susceptible/vulnerable subpopulation(s) were appropriate. The presented summaries suggested appropriately that individuals with asthma and potentially other preexisting lung diseases (e.g. COPD) are more likely to have an adverse outcome in response to short-term peak exposure to SO2 than individuals without preexisting disease.

Genetic background and age as susceptibility factors were also presented. While the REA appropriately indicated that limited data exist to reach a conclusion regarding the importance of genetic background as a susceptibility factor, the REA should include a statement indicating genetic susceptibility needs to be better characterized. Only one polymorphism has been evaluated for increased risk of susceptibility to SO2 effects (-308 *TNF* promoter SNP) and thus represents only a beginning. The revision of the draft REA provides an excellent opportunity to propose that a more thorough examination of genetic contribution is needed. The current evidence for genetic component of host responsivity to other criteria pollutants is strong (e.g. ozone), and it is likely that genetic variants will also be important in response to SO2.

Comments similar to the above can be made for differential responsivity attributable to age, although more studies currently exist that suggest age is an important susceptibility factor. Nonetheless, recommendations could be made for additional investigations to understand the relationship between age and response to 5-10 minute exposures to SO2, especially in the very young and elderly.

Dr. Timothy Larson

Comments on SOx REA – 1st Draft

1. We have evaluated SO_2 air quality throughout the United States, using all available 5-minute and 1-hour ambient monitoring data for years 1997 through 2007. To what extent are the air quality characterizations and analyses technically sound, clearly communicated, appropriately characterized, and relevant to the review of the primary SO_2 NAAQS?

The staff are to be commended for compiling and distilling this short term data. These analyses are relevant to the review of the primary NAAQS, given that there is strong evidence for effects from these short-term exposures above certain thresholds. These data are limited in geographical scope, but inclusion of the 5-minute maximum data as well as the continuous 5-minute data provides a reasonable data base.

2. To what extent are the properties of ambient SO_2 appropriately characterized, including ambient levels, spatial and temporal patterns, relationships between various averaging times, and the relationship between ambient SO_2 and human exposure?

The spatial variation of 24-hour and annual averages across the country on a large scale is well characterized. However, there are relatively few urban areas with multiple monitors and so it is difficult to assess intraurban spatial patterns based upon measurements. Therefore the reliance on plume models to infer the smaller scale variations is the only reasonable approach that is available. Those areas with multiple monitors been identified and given appropriate priority for inclusion in the larger modeling exercise.

The use of a pdf for the peak to mean ratios rather than applying a single value is appropriate. Defining a few different pdfs categorized according to SO_2 concentration and to proximity to major sources is a creative and useful approach that appears to converge to stable predictions in the final simulation.

The REA should clarify the major cause of the observed short-term peak concentrations, specifically whether these concentrations are correlated with peak emissions or unfavorable meteorology. As it stands, the implication is that meteorology is driving these peaks. The contribution to short-term ambient peaks from "upset" emission conditions at nearby major point sources as documented by continuous emission monitoring information should be clarified. If there is a correlation between the peaks and "upset" emissions, this would provide another way to assess the appropriate locations for analysis as discussed in question 3 below.

3. Twenty locations were selected for detailed analyses, using ambient SO_2 monitoring data for years 2002-2006. What are the views of the panel regarding the appropriateness of these locations, the time period of analysis, and the approach used to select them?

The three locations with relatively high 5-minute peaks are an obvious choice. The other 17 locations could have been chosen by any number of criteria. Choosing to limit the analysis to locations with multiple monitors in a given county is one reasonable approach aimed at capturing more spatial variation relative to a single monitor. Ranking the sites using the minimum adjustment factor (typically the one based on the 2nd highest 24 hour maximum) is reasonable.

However, one could also argue that some of these 17 additional locations could have been chosen based on potential for high downwind concentrations at locations other than the monitoring site. For example, using the emissions information in Table A-4 one can identify sources in Georgia, Kentucky, Minnesota, New York and Montana that have high emissions but whose locations are not included in the final list of 17. Some of these sources are somewhat isolated, but not all of them. Given that the exposure assessment in the REA predicts very few encounters with high 5-minute peak values at ground level, including some of these locations could alter the results. One approach is to do a simple screening level analysis based on plume impacts at all sites (e.g. Aerscreen) and then rank the locations.

4. In order to simulate just meeting either the current 24-hour or annual standards, staff adjusted SO_2 air quality levels for the years 2002-2006 upwards in all but one location. Ambient monitoring data in North Hampton County PA were above the 24-hour standard in the year 2006 and were therefore adjusted downward. To what extent is the approach taken technically sound, clearly communicated, and appropriately characterized?

The approach seems reasonable, given the lack of spatial information needed in order to include a space/time interaction (rather than the pure temporal adjustment based on one site applied equally to all sites). The approach is clearly communicated.

5. What are the views of the Panel regarding the adequacy of the assessment of uncertainty and variability?

Additional limitations include the fact that: 1) instances of building downwash of the plume is not being considered in the model (especially for older coal plants with relatively short stacks) and 2) that the effects of complex terrain are not being incorporated because the modeling locations chosen are not in such terrain. Monitors sited to capture the effects of building downwash or plume impaction on nearby, elevated terrain would measure higher peak hourly SO2 levels than if they

were located in flat terrain with unobstructed flow between the monitor and the stack, even if the emission rates are moderate. Those sited in the wake of buildings might also display different peak to mean ratios due to the different turbulence structure in this microenvironment.

Specific Comments:

p. 35 The pie chart in Figure 6-1 lists Hydroelectric Power Generation as a source of SO2. Are these emissions from facilities that combine both hydroelectric and coal-fired power plants? Hydroelectric plants by themselves do not emit SO2.

p. 42, line 17 should read "did not contain 5-minute measurements.."

pp 46-47 The monitor with the maximum number of 5-minute peaks in Figure 6-7 is actually located at the base of a ridge that runs between the Glover smelter stack and the receptor site. The site that is further away from this source is in open, flat terrain where the model presumably performs much better.

p 59 The bottom row of Figure 6-10 is presented as if it is a subset of the top row, yet the y-axis scales indicate the opposite is true.

p 125 line 20. Were any receptors located above stack height? Ditto for the results shown in Figure 7-3.

P143 Table 7-11 Other microenvironments could include the recirculating cavity induced by building downwash that is located next to a stack with less than GEP stack height and the elevated receptor on an isolated hill that is directly downwind at plume centerline height (the plume wrapping case under stable conditions aloft).

Dr. Kent E. Pinkerton

Comments on SOx REA – 1st Draft

Characterization of Health Risks (Chapters 3, 4, 5, 6, 7, 8, 9):

General Comments:

The first draft of the REA provides an excellent overview and extensive documentation that will be critical for the risk and exposure assessment plans in the review of the SO_2 national ambient air quality standard. The identification of sources for human exposure is important to clearly establish in order to better characterize personal exposure to ambient concentrations. County selection based on known SO_2 sources and archived SO_2 monitored data is excellent in providing substantive characterization for benchmark health risks for 5-minute peak SO_2 exposure. I fully agree with the designated at risk populations to SO_2 exposure and feel the human clinical studies are highly appropriate to form the basis for establishing the potential health effect benchmark values. The characterization of air quality and exposure analysis is impressive and presented in great detail.

Charge Question 4. To what extent is the assessment, interpretation, and presentation of initial risk characterization results technically sound, clearly communicated, and
appropriately characterized?

Response: It is my impression the assessment, interpretation, and presentation of initial risk characterization results are technically sound, clearly communicated, and highly reasonable in the manner it has been outlined and reported in this first draft. The approach for assessing exposure and risk associated with 5-minute peak SO_2 exposure is extremely reasonable and based on the findings of the controlled human exposure studies. Although 0.4 to 0.6 ppm SO_2 is being selected from these human clinical studies as the appropriate range to use in benchmark analyses associated with 5-minute peak SO_2 concentrations, it continues to be critical that 0.2 to 0.3 ppm peak SO_2 exposure also shows effects. Therefore, it is important to further justify the higher concentration of SO_2 exposure selected to use in this process. Also county selection for basing substantive characterization for benchmark health risks for 5-minute peak SO_2 exposure should be clarified to insure how representative each location is and how the sum of the findings will be applied across the country for risk and exposure in establishing a national standard.

Specific Comments:

Section 4.2.4 Decrement in lung function in the presence of respiratory symptoms (pages 22-23). For the study by Schwartz et al (1994), once co-pollutants were adjusted, was the SO_2 effect still significant? If so, please indicate. As stated, the effect is substantially reduced.

Although the staff has decided that it is not appropriate to use the epidemiological studies as the basis for a quantitative risk assessment, these studies continue to provide further validation of SO_2 exposure effects and should be given some consideration. It is good to see qualitative assessments of the epidemiology will be considered, but it would be good to specifically define how this qualitative assessment will be used.

Dr. Jonathan Samet

Comments on SOx REA – 1st Draft

General Comments:

This first draft REA for SO_2 provides extensive documentation for the plan on developing exposure profiles for the susceptible population. It still does not provide full details of the approach for assessing health risks, although the general framework is set out. The document would benefit if more overall structure were provided initially for the general approach that will be followed. In fact, it is not until chapter 8, which discusses the health risk assessment, that a general framework is offered for the risk and exposure assessment in Figure 8-1. It would be useful for readers if this figure were provided much earlier in the document. In fact, readers of the extensive chapters on assessment of concentration data and of exposure estimates would benefit from a better presentation of the overall structure of the risk assessment.

With regard to the characterization of health risks, my specific comments follow:

Charge Question 1

This question refers to the overall characterization of the health evidence for SO_2. The draft REA draws on the ISA in selecting the outcomes and exposure-response relationships to be used. The reliance on the clinical studies of persons with asthma is appropriate. There is a clear documentation of a causal association and the exposure-response relationship has been characterized with reasonable certainty. I am less certain as to the nature of the "qualitative" assessment that will be carried out using the epidemiological data (Charge Question 5). The positive risk estimates from the epidemiological studies selected will, of course, indicate an adverse effect. I did not find sufficient specificity on the approach and how the resulting information would be useful for assessing policy options.

The discussion of uncertainty and variability remains completely generic. At this point, while there is extensive discussion of these matters with regard to exposure, and a probabilistic approach is described for addressing uncertainty in health estimates, the overall approach in the risk characterization remains to be specified.

Specific Comments:

Chapter #-Page #	Line #	Comment
2	13-14	The concern with regard to misclassification arises in the

Chapter #-Page #	Line #	Comment
		context of hypothesis testing, and not necessarily with exposure assessment.
2-11	17	What is the distinction between "instantaneous" and "peak" exposure?
2-12	28-29	Is this uncertainty with regard to limited detection relevant to the discussion of peaks?
2-13	16	Replace "reliable" with "accurate"
2-13	22-23	This sentence is far too general and needs specificity.
2-13	23-24	While SO_2 levels may be difficult to measure at lower concentrations, they have little relevance to health.
2-14	3-5	The finding of low site-to-site correlations implies higher spatial variability.
3-16	20	Does this section on "age" refer to children and elderly persons with asthma?
3-17	11-14	The definition of "vulnerability" seems to have slipped from that in the ISA. Scenarios reviewed here refer to greater dose and not necessarily to a greater potential for exposure, the definition of vulnerability previously used.
4-22	16-18	In what way was the evidence found to be "most robust"? What was the criterion?
4-22	20	This comment concerning the epidemiological studies seems inconsistent with the view given that they do not address SO_2 alone.
4-23	6-9	The lag structure identified in this study seems quite inconsistent with the findings of the clinical studies. A comment is needed.
4-23	26-28	A change in the estimate with inclusion of additional variables in the model does not necessarily imply confounding.

Dr. Richard Schlesinger

Comments on SOx REA – 1st Draft

Overall, the document could use some general streamlining and editing to make it read "smoother." For example, in Chapter 2 it is not clear why information from earlier documents than the current ISA is repeated in some detail here (e.g., page 20 lines 1-13; p. 21 lines 12-18.) rather than presenting a summary of past studies. In addition, this document should use the details presented in the ISA to support the assessment approach used and not repeat details even of key studies (e.g., p. 20, l. 14-27). Chapter 7 is very di

p. 11, l. 20-21. This sentence is not necessarily totally true. While exposures are clearly likely in vicinity of source, SO2 is a regional pollutant as well and exposures may be in areas away from specific sources.

p. 17, l. 11-14. Here there is a mixing of susceptibility and vulnerability. Asthmatics are susceptible and people who work outside in general may be more vulnerable.

p. 19, l. 4 and l. 9. I think there is an error here in that the same terms are used in two places.

p. 116, Section 7.3.1. There needs to be better justification for use of data from Missouri when it was indicated that it was one of a few states that apparently had data that would allow for assessment of the modeling approach used.

Dr. Christian Seigneur

Comments on SOx REA – 1st Draft

Air Quality Information and Analyses (Chapter 6)

Overall, the air quality analysis is technically sound and appropriate for an SO_x risk and exposure assessment. My major comment pertains to Question 4: Some discussion on how an annual average air quality standards can be compared to 5-minute average values (see Section 5.2) is warranted. Alternatively, could EPA simply state that an analysis of the current annual NAAQS is inappropriate based on Table 5-3 of the ISA (see Section 4.1) since the presence or absence of any causal relationship cannot be inferred for any long-term exposure related effects (morbidity and mortality)? For example, Figure 6-30 shows that there is a fair amount of scatter between the number of exceedances of the 5-minute health benchmarks and the annual average SO_2 concentration. Then, only the short-term (24-hour average) standard would be analyzed.

Exposure Analysis

The exposure analysis chapters are clearly written and the overall technical approach is sound. The use of AERMOD for atmospheric dispersion modeling and APEX for population exposure estimates is appropriate. My major comment pertains to Question 2: The areas selected tend to focus on inland areas impacted by large stationary sources (coal-fired power plants, cement plants, chemical manufacturing plants, smelters). Thus, the potential impact of mobile sources is not directly addressed. As stationary sources undergo emission controls, the relative importance of some uncontrolled mobile sources (e.g., diesel-powered ships) may increase. Therefore, it would be worthwhile to model an area (in addition to Missouri) where ship emissions could have a significant impact on the population (e.g., Houston, TX or Los Angeles, CA).

Editorial comments:

p. 11, line 21: "principal" instead of "principles".

p. 12, line 15: add "coal-fired" before "electric generating units".

p. 14, line 15: "their" instead of "there".

p. 15, line 24: "attributable" instead of "attributible".

p. 19, the bullet on line 9 ("Short-term respiratory morbidity") under "inadequate to infer the presence or absence of a causal relationship" should be deleted since it is listed on line 5 as "sufficient to infer a causal relationship" (see Table 5-3 of ISA 2nd draft).

p. 29, line 21: “Canadian”.

p. 37, line 7: add “minute” after “continuous-5”.

Dr. Lianne Sheppard

Comments on SOx REA- 1st Draft

Air quality information and analyses (chapter 6):

This chapter relies exclusively on the monitoring data. The introductory section should be expanded to describe the purpose of the analyses.

While the analysis that relates the full 5-minute dataset to the 5-minute maximum dataset appears generally appropriate (there is an important exception noted below), the question of spatial representativeness is not considered outside of the universe of available monitors. What is the SO2 monitoring network supposed to represent? Is an unweighted summary of this network the best way to characterize 5-minute maxima?

Concern about the 5-minute dataset comparisons: The poor model fit at 2 monitors (see figure 6-7 p 47) needs much more careful investigation. Note that at the monitor with the highest number of measured exceedances, the number missed by the prediction exceeds the number of measured exceedances at any of the other monitors in the dataset.

There is something strange about 2004 in Figure 6-12 (p 61) that suggests some undocumented feature of the dataset that produces such a low normalized number of exceedances. The discussion on p. 58 mentions an Iron County Missouri monitor that ceased operation in 2003, but more needs to be done to determine if conclusions about trends reflect real phenomena or are merely features of the dataset that should not be generalized. This is one example of an aspect of the analysis that comes up several times in the chapter: it is important to be able to distinguish temporal trends in number of monitors in the network from downward trends in the concentration of SO2. Analyses need to be done to ensure that reductions in SO2 over time are real and not just an artifact of the change in the monitoring network. (For another example see the discussion on lines 8-12 p. 71.)

Many tables and figures need added clarification of titles, headings, or axis labels to ensure the reader doesn't interpret modeled or adjusted concentrations as though they are measured concentrations. The information may be in the caption, but it is easy to miss there. Examples include Table 6-12 (conc summary), Table 6-14 (conc summary), Figures 6-15, 6-23, 6-25, 6-30, 6-31, 6-32.

Add to Table 6-9 the number of monitors in each county and the number of neighborhood scale monitors.

The comparison of Figure 6-13 with 6-21 and 6-14 with 6-22 suggests much stronger correlation in the modeled than measured data, and monitors with much

higher number of exceedances at low values of annual average in the measured data than in the modeled data. These figures should be put on the same page and direct comparisons made. These comparisons suggest peaks may be underestimated, particularly for low annual average concentrations.

The uncertainty and variability discussion needs work. It is unclear what the analyzed monitors are supposed to represent spatially. This is relevant to a 6.51 conclusion (p100 l 29-30) since which monitors are included has significant impact on the generated results (even when the measurements themselves are of good quality). 6.5.4: There is almost certainly lack of spatial representation and uncertainty due to exclusion of monitors near local sources (102 l 28-30). Undefined spatial representation of monitors is one of the biggest challenges in making sense of the results presented in this chapter, and I would even discourage the assumption that data are representative of the locations analyzed (102 l 21-22) since the spatial scale is not defined in that comment and there is huge spatial variability in SO2 even within areas. 6.5.7: The section on the statistical model discusses the data, not the model. The reasons for exclusions need to be documented. The poor fit of the model in two locations is of concern and should be discussed.

Exposure analysis (chapters 2, 7):

Chapter 2 should look ahead to the use of estimates of exposure developed in chapter 7 for health risk analyses. Here are some questions:

- Is it surprising that there is poor site-to-site correlation of SO2 among monitors when these monitors are sited to capture local sources?
- If the number of 5-minute peak exposures to asthmatic individuals is as low as is estimated in chapter 7, is it worth continuing to the health analysis? This turns out to not be the right question, since chapter 8 uses the full range of the predicted SO2 exposure distribution and does not rely on exceedances alone. Can chapter 2 (or chapter 5) discuss the groundwork for the understanding of the exposure and risk assessment chapters?

Chapter 7 seems overall reasonable, with the exception of a few details discussed below. Assuming no changes, the key conclusion of the exposure analysis in this chapter is that in the modeled area, the number of potentially harmful exposures to at risk individuals for short 5-minute periods is low. Now with insight into the use of these predictions in chapter 8, the focus of chapter 7 needs to be expanded to also assess the entire distribution of predicted SO2 since that is used in the risk assessment. Finally, analysis should be done to align the estimates in chapter 6 with those produced in chapter 7 so the reader can understand why and how the two sets of estimates of peak exposures are different.

Concerns with the exposure model:

- p 142: The estimation of the additional 5-minute concentrations forces all the other values near the mean. This reduction in variability of the modeled data should effectively reduce the predicted number of 5-minute exceedances in any given hour. The analysis presented in Table 6-15 suggests this variance reduction will be too strong for the intended use of the modeled data.
- p 132: The comparison of measured data to the extremes of the distribution of modeled data appears to be a very weak test of the predictive capacity of the AERMOD model. Even so, Figure 7-4 suggests the predicted data don't capture most of the distribution of the measured data at that monitor, even if the upper tail is within bounds.

Response to charge questions:

- *To what extent is the assessment, interpretation, and presentation of the initial results of the exposure analysis technically sound, clearly communicated, and appropriately characterized?*

See comments above.

- *The draft risk and exposure assessment evaluates exposures in selected locations encompassing a variety of SO_2 emission source types in the state of Missouri; these areas were chosen as an initial case study since 1) air quality measurements indicated numerous exceedances of 5-minute benchmark values, 2) there are multiple stationary source emissions above 1,000 tons per year, and 3) there are numerous ambient monitors measuring 5-minute and 1-hour SO_2 concentrations.*

  The second draft may also evaluate exposures in the remainder of Missouri and also include areas of Pennsylvania, West Virginia, and other locations with large SO_2 emission sources. What are the views of the panel regarding the appropriateness of these proposed additional locations and on the approach used to select them?

The initial case study location selection is reasonable. I suggest only adding additional locations if the value-added can be defined. The results are a scenario based on a set of assumptions and local conditions. Are the local conditions as influential as the assumptions? It may be of more value to spend the modeling effort evaluating the sensitivity to the assumptions in the current location. Regardless of the choice, it needs to be stated several times in the document that the estimates apply to a limited population.

- *Do Panel members have comments on the appropriateness and/or relevance of the populations evaluated in the exposure assessment?*

These appear appropriate.

- *To what extent are the approaches taken to model SO_2 emission sources technically sound and clearly communicated?*

I defer some aspects of this response to my colleagues who know about implementation of AERMOD. With respect to evaluating predictions, the approach is better than the first draft NOx REA, but there is still room for improvement. The key consideration for quality of predictions for the purpose of this analysis is whether the predictions capture the variation observed in the real world. This includes both the peaks, which are the focus of the exposure estimation exercise, as well as the full exposure distribution, which is needed as input for the health risk assessment exercise. Figure 7-3 is a weak test of quality of predictions, but it is reassuring that the predicted maxima at all locations within 4 km are greater than the observed. Figure 7-4 is less reassuring, even for this weak test. It suggests most of the predicted distribution underestimates measured data, even though the peaks are reasonably comparable. I suggest also adding an evaluation of the AERMOD predictions at the receptor locations where there are monitors.

- *Human exposures were modeled using APEX to simulate the movement of individuals through different microenvironments. Do Panel members have comments on the microenvironments modeled?*

The modeling of movement of individuals appears appropriate. However the approach to estimating 5-minute peak concentrations will bias the results downwards. The distribution of number of peaks in an hour (Table 6-15) should be used instead of a procedure that sets all other observations in the hour at the mean. Peaks are likely to be correlated in time and this should also be factored into a revised algorithm.

Characterization of health risks (Chapters 3,4,5,7,8,9):

The summary of the health evidence from the ISA seems reasonable (chapters 3, 4). Chapter 5 is an introduction to the analyses in the rest of the document and could be used to lay out criteria for proceeding to later chapters and integrating interpretation across chapters (comment is particularly relevant to chapters 6-8).

The new chapter 8 results presented at the meeting look promising. It is notable that while chapter 7 focuses on benchmark values, chapter 8 uses a modeled dose-response function for the entire range of SO2 concentrations, so it targets different information in the exposure data. This suggests revisiting the modeling in chapter 7 to assess prediction of exposures below the peaks and adding the focus on the entire distribution to the chapter discussion. As a second point, the risk assessment will be sensitive to the assumed functional form of the exposure-response function. Since the data used to fit the function are so limited, the shape of this function and its behavior at the low end of the exposure distribution are highly uncertain. The 3-parameter logistic is particularly problematic. It fits a

limiting value parameter (α) which allows for a threshold in the population response for an exposure concentration that is below 100% (when $\alpha<1$). This is not justified scientifically. Since the shape of the function at low concentrations will be strongly impacted by the treatment of the curve at the upper end, this parameter must be justified. A sensitivity analysis of other functions should also be evaluated. I prefer locally defined functions (such as a linear spline or simple interpolation of points, at least for the FEV outcome) because their shape for low concentrations won't be as strongly informed by the fit at higher concentrations. I suggest at least one locally defined function be evaluated in sensitivity analyses.

The approach outlined in Chapter 9 may not help with understanding of population risk as extrapolated from time series study results. Time series relative risk models assume a log-linear concentration-response function. Given this parameterization of the risk model, summarization of the 98th and 99th percentiles of the concentration data should not lead to meaningful conclusions about the health effect estimates (unless perhaps the focus is on failure of the assumed log-linear dose-response model). I don't think the planned assessment of association of SO2 percentiles with the statistical significance of the epidemiological study results has any value. Along the lines of the suggested approach, I would prefer to see an analysis of the relationship between time series results in cities stratified by those using only monitors sited to capture background or community levels vs. those targeting local sources. Alternatively, I suggest EPA reconsider conducting a quantitative risk assessment for the epidemiological data based on time series study results. I do not think that SO2 falls on the "wrong" side of a bright line that distinguishes pollutants that are appropriate for quantitative risk assessment from those that aren't.

Concluding comments:

Please make sure the text reflects that numbers reported represent specific scenarios, such as a small subset of the population selected to be near SO2 sources. Thus these numbers can't be used directly to discuss estimates for the US population. An ongoing effort is needed to help readers properly interpret the numbers presented.

Page-specific comments:

- 19 l 5,9: short-term respiratory morbidity appears twice
- 37 l 20: Given the results in figure 6-2, it would be nice to replace the categorical parameterization with a linear one, if feasible.
- 40 l 3: Note how many were excluded by this QC procedure.
- 43: Improve the description in section 6.2.3.6 for clarity.
- 45 l 1: The underprediction at this monitor should be discussed in the context of its siting. (Looking at a google map representation, this monitor is near a source in a highly forested area. There is a hill that is taller than the stack that

forms a physical barrier that would affect measurements at the monitor location.) This topic comes up again (105 11-12). Further analysis of the site, such as discussed here, suggests why this site may be underpredicted. It also suggests features to assess in other sites in evaluation of predictions.

- 46 l 17: Is 1.02 the regression coefficient?
- 55 l 14: Table 6-5
- 58: Here is an example where the inference about SO2 trends over time is confused by the trends in monitoring. There needs to be an analysis that distinguishes these two features. (The current analysis shown in Figure 6-12 is a good attempt, but it doesn't tell the story clearly.) It may be better located in the ISA.
- 59 Figure 6-10: Add the number of monitors with no exceedances to the bottom of this figure. It may be necessary to add the total number of monitors to this figure as well, unless it can be shown elsewhere (e.g. in Figure 6-9). Same comment p 69 figure 6-16.
- 68 l 4: A COV of 0% does not seem reasonable. Some comment or appropriate data exclusion is warranted.
- 68 l 6: It does not appear useful to use data from a monitor with only 2 1-hour SO2 measurements. Also, clarify what "no below detection limit substitutions" means, particularly with respect to the mention that the dataset has 0 values.
- 71 l 4: The discussion regarding "this frequency would only apply" is unclear.
- 71 l 8-10: Is there evidence that the monitors with exceedances are being dropped? An analysis of existing data should be able to address this question so it will become unnecessary to speculate on this point.
- 111 l 5: Insert at the end of the sentence "when there are exceedances".
- 111 Table 6-15: Consider adding a row with 0 exceedances in the table for perspective (if it doesn't detract from the intent of the table).
- 111 l 20: Somewhere in this chapter there should be a discussion of the meaning of "peak" since in this chapter it is just high values, not maxima within a time frame (although the approach used effectively constrains the number of peaks to not exceed one per hour).
- 117 l 19: It will be worthwhile mentioning the non-point sources emissions used for SO2 in an appropriate place in the chapter.
- 118 l 5-6: Explain why the pairing of point sources with local meteorological stations was done.
- 118 l 13: Figure 7-1.
- 119 figure: why the random assignment to domains of receptors within multiple met station domains? Also better clarification of the symbols would help.
- 120 l 14: Sentence unclear.
- 124 l14-16: Comment on how much the stack locations moved and if impact is known, include that as well.

- 128 Table 7-5: Include an overall row with summations or average percentage, add "predicted" to the emissions for modeling, and merge cells so the total domain emissions column isn't confusing.
- 132 l 14-22: Since the exposure modeling in this chapter also needs to do a good job estimating the entire distribution for the risk assessment in Chapter 8, the underprediction of 95% of the distribution for AERMOD is troublesome. More should be done to refine the AERMOD prediction to better capture the entire distribution.
- 143 l 12,13: SO2
- 147 7.9: The "health risk characterization" focuses entirely on peaks which misleads the reader into believing this is the only purpose of the exposure modeling. This section should be revised to also address the input needed for chapter 8.
- 153 l 9-15: There is also the problem of constraining all remaining 5-minute concentrations to be equal to the mean.
- 168 19-20: The inference from this planned analysis should be discussed.
- Appendix A: I suggest adding the site classification(s).

Dr. Frank Speizer

Comments on first draft of Risk and Exposure Assessment SO2.
July 29, 2008

Discussion of Clinical studies: Page 15, line 28. This is taken from the ISA but there is an inconsistency in the ISA in that the actual study quoted showed 5-13% of subjects exposed to 0 .2 ppm for 5 -10 minutes had significant changes in sRaw and FEV1 respectively (see figure 4.1 in ISA). Thus to indicate that the effect level was "…as low as 0 .4-0.6 ppm …" is misleading. I certainly would not like to be in the group that dropped my FEV by 15%! This unfortunate statement is repeated throughout the next section and seems to set a quasi threshold for consideration of short term effects. This needs to be rethought with the idea of moving the minimal documented effect down from ">0.4ppm" to 0.2 ppm.

Discussion of the Epidemiological Short term studies. Although the studies are reasonably accurately reported they tend to ignore the phenomena indicated above. There are subgroups of individuals that as a class are more sensitive than others to SO2 and in most of the epi studies these subgroups are not considered. For example even among asthmatics, which as a group are believed to be more sensitive, there are individuals not sensitive and those that are. See above only 60% of asthmatics responded to 1ppm. Thus in the multicity studies or asthma ED studies there must be individuals who are not sensitive. As well as those that are extremely sensitive. So in reporting results as generally positive but not significant what is really being reported is positive results with wide confidence intervals generated by the misclassification of the "phenotype" of asthmatics that lumps together sensitive and non-sensitive subgroups. This needs to be discussed and if possible factored into the risk assessment calculations.

Section 5.2 From my first comment above it is clear that I believe that staff has chosen the wrong range for the benchmark analysis. They can go ahead and do 0.4-0.6ppm but they should also do the same analysis for 0.2-0.4ppm, since 13% of asthmatics are a big number.

Section 5.3: I agree with the plan to obtain more detailed SO2 air data from US and Canadian authors but isn't that totally impractical based upon the court ordered deadlines? I suspect that Staff will come back to us claiming they made the request maybe even got the data but it was too late to incorporate in analyses. Why not use existing network data to get the distributions out; without tying it specifically to data used by authors in studies that are now some years old?

Chapter 6 ambient air quality and benchmark health risks for 5 minute peak exposures.
This is an excellent start. One gets a reasonable "feel" for the available 5 minute average data. However, I would like to see similar plots for the exceedences of

>200ppb as well as these data at >400ppb. Specifically repeat paragraph bottom of page 56 along with figure 6-9 to 6-14 for >200ppb. I feel rather insistent that these calculations be done for lower levels and the justification is spelled out on page 112 section 6.5.9. Staff indicates as is the case that the studies reported for ethical as well as practical reasons were done on mild-moderate asthmatics. More severe asthmatics would be more susceptible that these mild asthmatics. Thus, with 13 %of such asthmatics having a 15% or greater drop in FEV1, at 200ppb it does not seem justified to start the risk assessment at 400 ppb where over 20% of mild asthmatics are responsive. This simply is the level of responsiveness that was measured in the clinical studies and to ignore it would be irresponsible.

Chapter 7: I note that in comparing table 7-2 to table 7-7 that although sites are designated by name and location in the former in the later they are all designated as Rural with the largest urban fraction being 17 and 19% and all others 5% or less. In addition all are air port locations. If this is the case these sights certainly do not represent population exposures. This could be a serious concern if there are regional sources located at these airport sites that impact the monitors. Some discussion, unless I missed it, should be presented on this issue.

The analysis suggested further in Chapter 7 that focuses on the Missouri sites does a good job of considering the model specification. This seems to work for these sites, perhaps because as stated on page 131 "all sources in Missouri are considered rural…" If this becomes the basis for the entire modeling of exposure something has to be done with longer range transport and more urban sites.

There is a discrepancy between tables 7-7 and 7-8 and table 7-9.* Two things in these tables don't make sense. Perhaps it is a decimal point placement. From the table ages 1 through 10 gives a total of about 1-2% asthmatics. It also indicates at the youngest ages Females outnumber males. Most studies I believe would say the opposite. Secondly in Table 9 and the text above suggest for these same ages about 10% of the children would be asthmatics (a more reasonable number). Need to adjust something.

Page 148, Section 7.9.2
This is were the selection of 0.4ppm vs 0.2ppm becomes important as the following tables show a very substantial differences in number of persons with exposure above a certain level.

In Chapter 8 again sets the stage with 0.4-0.6 ppm as the risk level. The discussion of uncertainty needs to include a section on what if the effect level is lower. (I clearly have indicated that I believe it is) Therefore the discussion might be turned around and after using the 0.2-0.4 numbers discussion the variability of response rather than uncertainty of findings.

* / This issue was discussed and resolved in the July 30-31, 2008 meeting. There was no discrepancy between the tables.

Chapter 9: I think it still might be worth considering the fact that many of the existing epi studies do show positive effects and some way of incorporating the fact that 15- 60% of asthmatics are responders to SO2 means that these overall effects that are not significant does not mean they are not positive. Therefore some risk assessment of the estimated responder populations might be worth calculating.
Look forward to seeing next draft.

Dr. George Thurston

Comments on SOx REA – 1st Draft

Air Quality Information and Analyses (Chapter 6):

1. We have evaluated SO_2 air quality throughout the United States, using all available 5-minute and 1-hour ambient monitoring data for years 1997 through 2007. To what extent are the air quality characterizations and analyses technically sound, clearly communicated, appropriately characterized, and relevant to the review of the primary SO_2 NAAQS?

RESPONSE: The data appear to be the best available for the analyses attempted, but they need to be subdivided by monitor type, especially source-oriented vs. not near a major SO_2 source. In addition, the peak-to-mean ratio model (Equation 6-1) seems overly simplistic, in that it does not implicitly address the variability in the COV. Instead, the sites are placed in "bins" according to their COV, which means that a range of COV's are handled similarly. Instead, it would seem that fitting another model term (dependent on the COV) would be a more appropriate approach, and might avoid the outliers found when testing the bin model (pages 45-46). In addition, it is not clear to me that the test of goodness of fit is independent of the original fit…is it? Or is the EPA just testing the model on the fit derived from the same data? The best situation is to develop the model on one set of data, and test it on another separate set of data. Please clarify which data were used to fit the model, and which were used to test the fitted model.

2. To what extent are the properties of ambient SO_2 appropriately characterized, including ambient levels, spatial and temporal patterns, relationships between various averaging times, and the relationship between ambient SO_2 and human exposure?

RESPONSE: This seems to have been accomplished the best that can be done with the data available for the purpose. However, it would be helpful to sub-characterize these data as a function of site type (i.e., source oriented vs. other categories), so as to better understand how the populations most at risk (i.e., near major SO_2 sources) differ from people located elsewhere.

3. Twenty locations were selected for detailed analyses, using ambient SO_2 monitoring data for years 2002-2006. What are the views of the panel regarding the appropriateness of these locations, the time period of analysis, and the approach used to select them?

RESPONSE: Acknowledging that data limitations do exist, it still seems to me that this analysis should focus on areas where violations are most likely, i.e., in counties where major point sources exist, such as in and around Jefferson County, Ohio. Therefore, I think an additional source-oriented criteria should be added to focus the analyses more on areas where the problem of concern here (i.e., high

peak impacts) is most relevant. Moreover, the fact that the Source-oriented Caribou, ID site is poorly fit (pg. 70) may in fact indicate that the model is not performing well in situations of greatest interest in this analysis. Finally, the benchmarks employed are too high, and should be lowered, as even chamber studies of pure PM have exhibited effects down to 200ppb (e.g., see Figure 4-1 og the REA), and the animal toxicology indicates that effects are seen at much lower levels when particles are co-present with the SO_2, in agreement with the epidemiology showing associations at ambient-level short-term SO_2 (e.g., Peel et al, 2005).

Thus, the last sentence on page 112 should instead read something more like: "Therefore, the potential health effect benchmarks based on these clinical studies likely underestimate risks in the general population because people in the general population with greater susceptibility are considered, and the exacerbating effects of co-present ambient particulate matter are also not considered in such a limited analysis of risk.

4. In order to simulate just meeting either the current 24-hour or annual standards, staff adjusted SO_2 air quality levels for the years 2002-2006 upwards in all but one location. Ambient monitoring data in North Hampton County PA were above the 24-hour standard in the year 2006 and were therefore adjusted downward. To what extent is the approach taken technically sound, clearly communicated, and appropriately characterized?

RESPONSE: Yes, this seems a reasonable approach to estimation.

5. What are the views of the Panel regarding the adequacy of the assessment of uncertainty and variability?

RESPONSE: Again, it is not clear to me that the uncertainty analysis considers a dataset distinct from the data used to develop the model in the first place (e.g., in the accuracy estimation on page 105). Please clarify this in the text. Moreover, I feel that a quantitative risk assessment based upon the SO2 epidemiological studies (e.g. of respiratory ED visits) could provide a useful input to the uncertainty analysis. Since the number of people ending up at the Emergency Department should be a small subset of the number of people experiencing bronco-constriction, a quantitative estimate of the SO_2 associated ED visits would provide a check on whether the number of people affected are being underestimated. Such an analysis would be a useful addition to the assessment of uncertainty.

Exposure Analysis (Chapters 2, 7):
1. To what extent is the assessment, interpretation, and presentation of the initial results of the exposure analysis technically sound, clearly communicated, and appropriately characterized?

RESPONSE: None on this aspect.

2. The draft risk and exposure assessment evaluates exposures in selected locations encompassing a variety of SO_2 emission source types in the state of Missouri; these areas were chosen as an initial case study since 1) air quality measurements indicated numerous exceedances of 5-minute benchmark values, 2) there are multiple stationary source emissions above 1,000 tons per year, and 3) there are numerous ambient monitors measuring 5-minute and 1-hour SO_2 concentrations. The second draft may also evaluate exposures in the remainder of Missouri and also include areas of Pennsylvania, West Virginia, and other locations with large SO_2 emission sources. What are the views of the panel regarding the appropriateness of these proposed additional locations and on the approach used to select them?

RESPONSE: I'd like to see more analyses of locations near major SO_2 sources, like power plants, in counties such as Jefferson County, OH, and surrounding counties.

3. Do Panel members have comments on the appropriateness and/or relevance of the populations evaluated in the exposure assessment?

RESPONSE: The populations considered (i.e., asthmatics) may be too narrow for the standard-setting process, which would lead to small estimated numbers of people affected. Rather than clinical studies of subset populations, and only to pure SO_2, the application of epidemiology-based risk factors would provide a greater relevance to the general population. At a minimum, risk estimates (e.g. of ED visits) should be conducted based on epidemiology in order to determime if the population considered is plausible. The number of SO_2 induced respiratory ED visits should be a subset of the population at risk, so if the epidemiology gives larger numbers, that would provide a test as to whether the full population at risk has been considered, or not.

As to the asthmatic populations to be considered, it may be important to consider that the people in the general public who have regular anti-inlammatory medications prescribed by their physician may represent an especially affected subpopulation, especially on a day when they have not taken their prescribed medication.

4. To what extent are the approaches taken to model SO_2 emission sources technically sound and clearly communicated?

RESPONSE: Appears to be state-of-the-art and well explained. The dependence on likely incomplete source emissions inventories is a potential weakness.

5. Human exposures were modeled using APEX to simulate the movement of individuals through different microenvironments. Do Panel members have comments on the microenvironments modeled?

RESPONSE: No.

Characterization of Health Risks (Chapters 3, 4, 5, 6, 7, 8, 9):
1. What are the views of the Panel on the overall characterization of the health evidence for SO_2? Is this presentation clear and appropriately balanced?

RESPONSE: Section 8.2.3 needs to clearly point out that these studies are for pure SO_2 only, and do not fully represent conditions in the real world, as the SO_2 interactions with PM (that is always present in the real world) are not considered. This may well lower the levels at which the symptoms noted can be experienced. The mechanism for this is likely that the particles provide a vector for sulfur oxides to be transported deeper into the lung in solution and as reactant products.

Indeed, Chen *et al.* (1992) have revealed that approximately 10 times as much pure sulfuric acid (H_2SO_4) is required to give the same lung airway hyper-sensitivity effects in guinea pigs as when the acid aerosol is present as a surface coating on a particle (200 *u*g/m^3 H_2SO_4 mist vs. 20 *u*g/m^3 H_2SO_4 when surface coated on a particle). (Chen LC, Miller PD, Amdur MO, Gordon T. (1992). Airway hyperresponsiveness in guinea pigs exposed to acid-coated ultrafine particles. J Toxicol Environ Health. Mar;35(3):165-74.) Sulfuric acid is one potential surface reactant product of SO_2 and particle-surface reactions. Thus, it might well be possible that, in the real world where particles are always co-present, the acute effects noted with pure SO_2 at 200 ppb may well be experienced at much lower SO_2 exposure concentrations, and this should be considered here and throughout this document. Moreover, it might be argued by some that this effect is covered by the PM standard, but the very acute effects considered here are associated with SO_2, and, also, there is no one-hour or 5-minute PM standard, so even though PM co-presence is apparently involved in exacerbating the impact of SO_2, the effects under consideration here are something very distinct from the longer averaging time PM effects controlled by that standard, and must be considered in this document and SO_2 standard-setting process.

2. The characterization of health risks focuses on potential health benchmark values identified from the experimental SO_2 human exposure literature on lung function with accompanying respiratory symptoms. What are the views of the Panel on using potential health benchmarks from this literature to characterize health risks?

RESPONSE: Such controlled clinical studies of pure compounds are very important for proof of concept and for evaluating biological plausibility, but not for risk assessment as proposed here. Epidemiological studies should be applied for that process, as they consider real people in real world situations.

3. Do panel members have comments on the range of potential health effects

benchmark values chosen to characterize risks associated with 5-minute SO_2 exposures?

RESPONSE: The benchmarks selected are too high. First, there are effects documented in the pure SO_2 clinical exposure studies at levels at least down to 200 ppb. Second, the exacerbating effects of co-exposure to PM on the health impacts of SO_2 exposure in the real world is ignored: it is plausible that co-exposure to PM will cause these effects at much lower levels than indicated by the clinical exposures to pure SO_2 alone. Finally, the epidemiology concur with this point in that they show associations that the ISA finds sufficient to infer a causal relationship (see page 19 of the REA).

4. To what extent is the assessment, interpretation, and presentation of initial risk characterization results technically sound, clearly communicated, and appropriately characterized?

RESPONSE: I find the assessments based on the clinical studies to be insufficient alone, as the full extent of effects in the real world on a wider distribution of the population cannot be fully incorporated using this approach. Epidemiology should also be applied in this process, if only as a quality control check on the plausibility of the size of populations affected. As noted above, a comparison of clinical study-based estimates vs. epidemiology-based estimates would provide at least a check on the plausibility of the population estimated as affected based on the clinical studies that are based on pure SO_2 alone.

5. The epidemiology literature will be used to qualitatively characterize SO_2-related health risks for health outcomes such as respiratory symptoms and emergency department visits and hospital admissions for respiratory-related causes. However, staff has judged that it is not appropriate to use the available SO_2 epidemiological studies as the basis for a quantitative risk assessment in this review. Do panel members have comments on this judgment and/or on the rationale presented to support it?

RESPONSE: While this is a worthwhile analysis to conduct, I feel strongly that the epidemiological studies of SO_2 can and should also be used to conduct a quantitative risk assessment, if only as a check on the clinically-based estimates. Furthermore, I don't think that looking at the correlation of percentile SO_2 concentrations vs. statistical significance is a worthwhile, or very meaningful, exercise. Too many other variables (such as power) enter into the determination of statistical significance for this to be a meaningful exercise. The EPA should move forward with a quantitative risk assessment based on the epidemiological studies available, albeit noting the uncertainties and limitations, in order to provide a fuller and more relevant risk assessment than allowed via relying only upon the clinical studies-benchmark approach for quantitation, as they propose in this draft document.

Dr. James Ultman

SOx REA Comments

General Comments on the Document

The first draft of the REA clearly puts forth the susceptible population (i.e. asthmatics) and health effects (i.e., clinically-observed symptoms and lung function decrements) that will be the focus of the health risk assessment. It is also is evident that this REA will extend previous assessments by a detailed analysis of the consequences of short-term and peak exposures under alternative forms and levels of the NAAQS.

Chapter 6.

General

Because of the limited number of monitoring data on peak exposures, staff has developed an imaginative but previously-unvalidated stoichastic method to extrapolate from short-term hourly exposure data to peak exposure concentrations. The rationale for the method is that "..the temporal and spatial pattern in SO2 source emissions is influenced by the type(s) of sources and its operating conditions and that this emission pattern(s) will be reflected in the ambient SO2 concentration distribution measured at the monitor." Based on this rationale, the coefficient of variation (COV) of 1-hour exposure measurements is used as a predictor of the peak-to-mean ratio (PMR) of the hourly measurements.

The selection of COV as a predictor variable is justified by analyses of the data from 98 monitors where co-localized peak and hourly averaged SO2 concentrations. These analyses include: the linear correlation of the COV's of 5-minute samples with the COV's of 1-hour samples (figure 6-2); the convergence of the predicted PMR values to the measured PMR values (figure 6-5); and a comparison of the mean predicted PMR value to the measured PMR at each monitor.

The latter analysis is presented as a test of the "accuracy" of the PMR estimation method. Since the measured values used to evaluate the method is the same data set used to obtain the cumulative distribution functions (CDF) used in the simulations, this analysis does not validate the method. It would have been better, in theory, to divide the 98 monitors into two subsets—one subset for determining the CDF and another for validating the method.

Even after reading appendix A, I find the details for the many algebraic computations performed in this chapter hard to follow (e.g., see lines 6-12). Such computations would be more transparent if they were presented as equations, or even better, supported by idealized graphs that showed how a (hypothetical)

concentration-time trace from a monitor was averaged over 5-minute and 1-hour intervals that were then averaged together, etc.

I also find it hard to follow the progression of analyses in sections 6.4 and 6.5. It appears that the "as is" analysis of exceedances above the health effects benchmarks is obtained from the full 98-monitor data set, whereas the "just meeting the current standard" analysis of exceedances is obtained from the 20-county data set. If this is indeed the case, then is would be inconsistent to compare the two analyses. To avoid such confusion, the chapter would benefit from a more informative introduction, either at the beginning of the chapter or at the beginning of the major sections.

Specific Comments

Page; lines

33; 11 Spatial siting of monitors should, in principle, impact both horizontal as well as vertical distances from point sources. Are the distributions of vertical distances of the 98 monitors upon which the PMR method is based similar to the vertical distances at which all 1-hour monitors are placed?

36; **12** This "model" equation gives the impression that PMR is a parameter. In fact, PMR depends on $C_{1\text{-hour}}$. It might be better to write the equation as a definition of PMR.

39; **8** Does the Thompson reference provide validatation the stoichastic approach used in the current document?

42; 3-5 I don't see why these results are "consistent" with each other. Perhaps, more explanation is needed.

46; 11 I wouldn't say that the table entries exhibit "good agreement."

Responses to the Charge Questions – Air Quality Information and Analysis

1. The clarity and flow of the many analyses in this chapter could be significantly improved. The mechanics of a particular analysis are not always clear. Moreover, the relationship among the many analyses is hard to follow. With respect to the technical aspects of the chapter, I feel that additional thought needs to be given to validating the PMR estimation procedure.
2. There is an abundance of basic numerical information in the chapter, but at some point, it needs to be distilled into a set of more easily appreciated observations and conclusions.
3. No comment.

4. It was not clear to me, from the contents of this chapter, how the roll-up factors determined in 20 selected counties will be applied to the exposure and health risk assessment on a national level.
5. The primary source of uncertainty is the lack of validation of the PMR methodology.

Dr. Ronald Wyzga

Comments on SOx REA – 1st Draft

Overall Comments: It would have been helpful to have had more time to review this document. It is a very lengthy and complex document. Given the available time to review it, my review is at best cursory.

By and large, I find the approach taken in this document to be reasonable. The assessment focuses upon short exposures to asthmatics, which I believe to be the key issue for SO2.

Charge Questions for Exposure Analysis:

1. To what extent is the assessment, interpretation, and presentation of the initial results of the exposure analysis technically sound, clearly communicated, and appropriately characterized?

My review is cursory, but at first glance it appears to be technically sound and appropriately characterized.

2. The draft risk and exposure assessment evaluates exposures in selected locations encompassing a variety of SO2 emission source types in the state of Missouri: these areas were chosen as an initial case study since 1) air quality measurements indicated numerous exceedances of 5-minute benchmark values, 2) there are numerous ambient monitors measuring 5-minute and 1-hour SO2 concentrations. The second draft may also evaluate exposures in the remainder of Missouri and also include areas of Pennsylvania, West Virginia, and other locations with large SO2 emission sources. What are the views of the panel regarding the appropriateness of these proposed additional locations and on the approach used to select them?

I agree that attention should be given to those areas where there are exceedances presently and where there are major SO2 sources. Given its size, particular attention should be given to Allegheny County (Pittsburgh), Pa. If there is a tradeoff in resources between extent of detail in estimating exposures and the number of areas studied, I would favor emphasis on the former. I think Missouri, Pennsylvania, West Virginia, and possibly Ohio would provide a good understanding of the risks in states where exposures are above average. If there are any remaining instances of high exposures associated with smelter operations, these might be considered as well.

3. Do Panel members have comments on the appropriateness and/or relevance of the populations evaluated in the exposure assessment?

I would agree with the focus on asthmatics. I note the apparent discontinuities in some asthma prevalence rates in Table 7-7; can these be verified? It would be useful to obtain data on the relative number of asthmatics who are routinely medicated as this group does not appear to respond to peak SO2 exposures.

4. To what extent are the approaches taken to model SO2 emission sources technically sound and clearly communicated?

Given my limited expertise in the use of air quality models, I leave it to my colleagues to judge this issue.

5. Human exposures were modeled using APEX to simulate the movement of individuals through difference microenvironments. Do Panel members have comments on the microenvironments modeled?

The APEX model is well-suited for exposure analyses to be undertaken here. My only question about microenvironments is whether roadside exposures should be considered. I have been involved in some studies which suggest that meaningful exposures to SO2 can occur from sulfur-containing diesel fuels (which are being phased out); if this is correct, near-roadway exposures, could be higher. On the other hand since this source of SO2 is being curtailed significantly, it could be fruitless to consider this source in future regulatory scenarios.

Specific comments:

p. 13, ll. 26-28: Should special note be made about the amount of time spent indoors as indoor exposures are negligible except in the rare cases where there are indoor sources.

p. 15, l. 22: insert “exercising” before “asthmatics”.
11. 28 and follows: should a comment be made that exercising asthmatics who are medicated do not appear to respond to SO2 in human clinical studies.

p. 34, l. 8: “hydroelectric”???
l. 19: replace “is” with “are”.

p. 36, l. 12: Is this equation too simple? Do we need to consider wind direction?

p. 37, l. 16: Which 6 states?

p. 42, ll. 18-25: I wonder if there are better ways to do this, by considering the proximity of monitors to sources and/or considering such factors as wind speed. I would be interested in more details about the distribution of COVs as well.

p. 47, ll-17-18: In general I worry about the communication of this scenario; it can be very misleading when so few areas exceed the current standard. I hope this scenario is well-caveated.

p. 54: I wonder if it useful to consider adjustments based upon the annual average concentration given the uncertainty associated with the relationship between short-term concentrations and the annual average.

p. 55, ll. 12-13: Is there also a statistically significant trend?

p. 71, l. 3: "than"

p. 135, Table 7-7: There are some curious discontinuities in the prevalence rates by age, especially for males; see the differences between age 3 and 4, 4 and 5, and 16 and 17. Are these numbers correct?

p. 145, ll. 14-17: Are there any data to suggest some consideration of near roadway exposures? Sulfur in diesel fuel may have influenced such exposures in the past.

p. 158, l. 4: "Introduction"

p. 165: Medication use could be another category of uncertainty.