UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON D.C. 20460

OFFICE OF THE ADMINISTRATOR
SCIENCE ADVISORY BOARD

January 9, 2008

EPA-CASAC-08-006

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

Subject: Clean Air Scientific Advisory Committee's (CASAC) Consultation on EPA's *Sulfur Dioxide Health Assessment Plan: Scope and Methods for Exposure and Risk Assessment* (November 2007 Draft)

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, met on December 6, 2007 for a consultation on EPA's *Sulfur Dioxide Health Assessment Plan: Scope and Methods for Exposure and Risk Assessment (*November 2007 Draft*)*. The CASAC uses a consultation as a mechanism for individual technical experts to provide comments to guide the Agency on technical issues early in the development of a document, before the first draft is ready for peer review. Panel members offered oral comments at the meeting as well as written comments (attached to this letter). This CASAC consultation, like all CASAC consultations, was conducted under the requirements of the Federal Advisory Committee Act, which include advance notice of the public meeting in the Federal Register.

There will be no formal report from the CASAC as a result of this consultation, nor do we expect any formal response from the Agency. CASAC offers the attached individual comments for the Agency to consider as it moves forward with his Health Assessment for Sulfur Dioxide.

We look forward to conducting a peer review of the first draft for the Exposure and Risk Assessment document as part of the CASAC's continuing role in the National Ambient Air Quality Standard review.

Sincerely,

/Signed/

Dr. Rogene Henderson, Chair
Clean Air Scientific Advisory Committee

Attachments

Attachment A: Roster of CASAC Sulfur Oxides Primary NAAQS Review Panel
Attachment B: Compilation of Individual Panel Member Comments on EPA's *Sulfur Dioxide Health Assessment Plan: Scope and Methods for Exposure and Risk Assessment* (Draft, November 2007)

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

CASAC MEMBERS

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

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

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

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

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

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

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

PANEL MEMBERS

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

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

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

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

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

Dr. Steven Kleeberger, Professor, Lab Chief, Laboratory of Respiratory Biology, National

Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, NC

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

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

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

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

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

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

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

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

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

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

SCIENCE ADVISORY BOARD STAFF

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

Comments from CASAC Sulfur Oxides Primary NAAQS Review Panel on EPA's Draft Sulfur Dioxide Health Assessment Plan: Scope and Methods for Exposure and Risk Analysis (November 2007)

Comments from Dr. Cowling

My comments are organized below in response to each of the several Charge Questions posed in Karen Martin's November 2007 transmittal letter for Lydia Wegman to Holly Stallworth.

Air Quality Considerations:

1. Based on the low estimated contribution of policy-relevant background SO_2 to overall ambient SO_2 levels, staff is considering a proportional (i.e., linear) approach to adjusting air quality to simulate just meeting potential alternative SO_2 standards that are below recent air quality concentrations. Do the Panel members have comments on adopting a proportional approach to simulate just meeting more stringent alternative air quality standards?

Such a proportional approach seems very sensible to me

2. Recognizing that current ambient air quality concentrations are lower than the current standards, the draft Health Assessment Plan discusses two alternative approaches to simulating ambient SO_2 levels associated with just meeting the current SO_2 standards: use of historical air quality data (e.g., possibly pre-2000) when ambient levels were at or above the current standards, or use of a proportional (i.e., linear) approach to adjust SO_2 levels upward. Do the Panel members have advice or comments on these two alternative approaches to simulating air quality just meeting the current SO_2 standards?

Being a student of history I would favor using historical data to simulate air quality parameters that just meet the current standards.

Exposure Analysis:

1. In considering the exposure analysis broadly:

a. Do Panel members have any comments on the general structure and overall two-tier approach that staff plans to use for the exposure analysis? Are the criteria that staff plans to use for deciding whether to conduct a Tier II analysis clear and appropriate?

The description of the two-tier approach is outstandingly clear as presented in this document. Unfortunately, however, I have no personal experience on which to base an informed judgment about the issue of appropriateness of the two-tier approach.

b. Have the most important factors influencing exposure to SO_2 been clearly accounted for and described?

Yes. It appears to me that the important factors influencing exposure have been accounted for and have been described clearly.

c. The draft plan describes the basis for and selection of population groups of interest (i.e., children, asthmatics (children and adults), and the elderly) for which SO_2 exposure estimates are to be developed. Do Panel members generally agree with the groups of interest identified in the draft plan?

I certainly agree with the population groups (of interest or concern) that have been identified in the draft plan.

2. <u>In considering the Tier I exposure assessment</u>:

a. Do Panel members agree that a statistical model using available ambient 5-minute monitoring data is appropriate for estimating expected exceedances of very short-term (5-minute) potential health effect benchmarks?

I have no experience on which to base an informed judgment in response to this question.

b. Do Panel members agree with the approach of applying a statistical model to estimate 5-minute concentration exceedances at monitoring locations where only 1-hour monitoring was performed for evaluating the extent of 5-minute peaks associated with meeting alternative standards with longer averaging times?

I presume there is an adequate body of measurement data where both 5-minute and 1-hour measurements have been made in various locations across this country, and that the correlations between these parallel measurements can provide an adequate basis for developing a statistical model of reasonable reliability. If such parallel data sets are not available, or if correlations between 5-minte and 1-hour data are highly variable, however, it seems risky to use a statistical approach of indeterminate reliability. See Checklist question 3 in the "Guidelines for Formulation of Statements of Scientific Findings to be Used for Policy Purposes:"

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

3. <u>In considering a potential Tier II exposure assessment</u>:

a. Do Panel members agree with the combined emissions/dispersion modeling approach to estimate short-term (hourly) SO_2 concentrations in close proximity to SO_2 emission sources?

I have no experience on which to base an informed judgment in response to this question.

b. Do Panel members have comments or advice regarding the described binning of sources and development of prototype stacks/facilities?

I have no experience on which to base an informed judgment in response to this question.

c. Do Panel members agree with the approach using peak-to-mean ratio cumulative density functions (PMR CDFs) to estimate very short-term peak concentrations from the 1-hour modeled concentrations?

I have no experience on which to base an informed judgment in response to this question.

d. Do Panel members generally agree that the approach described using APEX is reasonable and appropriate to estimate the occurrence of very short-term (5 minute) SO_2 peak exposures?

I have no experience on which to base an informed judgment in response to this question.

4. Do Panel members have any comments or advice regarding the general approach to addressing uncertainty and variability in each Tier of the exposure assessment as described in the draft plan?

I find the description of the general approach to addressing uncertainty and variability in each Tier of the exposure assessment very clear as presented in the draft plan.

<u>Health Risk Assessment</u>:

1. Do Panel members have any comments on the general structure and overall three-tier approach that staff plans to use for the risk assessment? Are the criteria that staff plans to use for deciding whether to conduct a Tier III risk assessment clear and appropriate?

The description of the three-tier approach is outstandingly clear as presented in this document. Unfortunately, however, I have no experience on which to base an informed judgment about the issue of appropriateness of the three-tier approach.

2. In considering the Tier I risk assessment:

a. Do Panel members agree with the approach of having a qualitative assessment of health endpoints to identify which are likely candidates for a more sophisticated and quantitative tier of assessment?

Although it seems reasonable that a qualitative assessment of health endpoints might be used to identify likely candidates for a more sophisticated and quantitative tier of assessment. As indicated earlier, however, I have no personal experience on which to base an informed judgment about the use of qualitative assessments in making choices about quantitative tier assessments.

b. Do Panel members agree with our initial observation that controlled human exposure studies demonstrate strong evidence for bronchoconstriction in exercising asthmatics following 5-10 minutes SO_2 exposure?

I have no experience on which to base an informed judgment in response to this question.

c. Do Panel members agree with staff's initial observation that the strongest epidemiologic evidence is for respiratory symptoms in asthmatic children and respiratory-related hospital admissions and respiratory-related emergency department visits in asthmatics and others with respiratory conditions?

I have no experience on which to base an informed judgment in response to this question.

3. In considering the Tier II risk assessment:

a. In general, are staff plans to use potential health effect benchmarks to address respiratory effects demonstrated in exercising asthmatics in controlled human exposure studies clear and appropriate?

I have no experience on which to base an informed judgment in response to this question.

b. Do Panel members generally agree with the tentatively identified potential health effect benchmark of 0.5 to 0.6 ppm for exercising asthmatics following 5-10 minutes SO_2 exposure?

I have no experience on which to base an informed judgment in response to this question.

c. Do Panel members generally agree with the staff's approach of focusing on areas around major sources of SO_2 with respect to concerns about 5-10 minute peak exposures related to the respiratory effects observed in controlled human exposure studies?

Yes, this approach seems very reasonable to me.

d. Do Panel members generally agree with staff's approach of focusing on urban areas with respect to concerns about 1- and 24-hr and annual SO_2 concentrations related to respiratory effects observed in epidemiologic studies?

I have some misgivings about focusing so strongly on SO_2 concentrations in urban areas that people (including both susceptible and vulnerable populations in other regions with somewhat higher exposures (such as the Pacific Northwest, Hawaii and Alaska) may be short-changed in the planned assessment processes.

e. Do Panel members have any comments or advice with respect to staff's approach of gathering additional information to characterize the SO_2 ambient air quality that existed at the time various key U.S. and Canadian studies addressing respiratory effects were conducted to see if the concentration-response relationships observed in these epidemiologic studies are related to particular SO_2 levels and associated averaging times, geographic location and/or season, and the inclusion of various copollutants?

I have no experience on which to base an informed judgment in response to this question.

4. In considering a potential Tier III risk assessment:

a. Do Panel members generally agree that there is insufficient information to develop credible exposure-response relationships for use in a quantitative risk assessment based on the controlled human exposure evidence?

I have no experience on which to base an informed judgment in response to this question.

b. Do Panel members have any comments or advice with respect to the general approach or specific factors to be considered in deciding whether or not to proceed to a Tier III quantitative risk assessment for the respiratory-related health endpoints based on epidemiologic evidence discussed in the draft plan?

I have no experience on which to base an informed judgment in response to this question.

5. Do Panel members have any comments or advice with respect to the general approach to addressing uncertainty and variability in each Tier of the risk assessment as described in the draft plan?

I am very impressed with the clarity of presentation of the general approaches described in this draft plan for addressing uncertainty and variability in each Tier of the proposed risk assessment! At the same time, however, I can only trust that skill in development of written descriptions is not a cover for lack of skill within the assessment team for drawing of appropriate scientific inferences from analysis of available data and information!

Comments from Dr. Crawford-Brown

My overall impression of the document is that it presents a reasonable path forward on assessing risk in a topical area (sulfur dioxides) where the data are somewhat sparse. However, it took me several readings to piece together the structure of the assessment process due to poor organization of the document. The key point of confusion was that the authors describe a process that is different for effects that have been studied through clinical trials and ones that have been studied through epidemiological results. This produces in two different ways of characterizing risk for these two categories of effects, and two different levels of detail in the characterization. This separation – or at least the basis for it - is not made evident, however, until late in the document, and so the reader is left partially confused in the first two thirds. I kept sensing that there were two streams of thought and assessment at play, but never had a concrete statement of that until late in the document. The writing overall needs to be improved.

This improvement is needed both to clarify the issue that there are two kinds of assessments (for the two kinds of data) and to explain how precisely the epidemiological data are to be assessed. Despite several readings, I cannot understand what they intend to do with the epidemiological results. For the clinical results, they clearly intend a modified form of hazard quotient or margin of exposure. That will result in a quantitative measure of at least hazard. But in the case of the epidemiological results, the authors talk repeatedly of "looking for patterns" in the concentration-response data, with no indication of what they mean by "patterns". I suppose they might mean looking for changes in the slope factor or some other risk summary as one moves across studies of concentration-response in different geographical settings and populations, but there is never a succinct statement as to what they mean by a "pattern" so I remain unsure. And it is not at all clear what they intend to do with such patterns, or even what the measure of hazard or risk will be at the end of the day. I see no reason why it cannot be a benchmark health effect as in the clinical trials, with a similar calculation of hazard quotient or margin of exposure. This part of the document needs to be significantly improved.

I have a series of more specific comments:

1. The first full paragraph on page 2 appears to indicate that the focus will be on short-term exposures that take place in close proximity to local source emissions. The rest of the document, however, does not appear to narrow the focus so tightly. This confusion needs to be clarified.

2. The tiering system is OK but still somewhat confusing to read. The first problem is that it is not clear what specific results from a first tier would send the assessors to the second tier (or what results would prevent them from going there). Some VERY loose criteria are mentioned (e.g. on Page 36), but these are quite generic and the real question is what kinds of answers to these issues would constitute staying in a tier or advancing.

The second problem is that there are two tiers for exposure and three for risk. I suppose the authors intend that the exposure assessment could proceed to a different tier than the risk assessment (yielding 6 cells in a 2 x 3 table), but that seems to me an unwarranted approach.

Better to have just two tiers that apply to the entire process. I see no merit to having, for example, a tier 2 exposure assessment and then a tier 1 risk assessment. The tier 2 exposure assessment would contain a level of detail that could not be met by the concentration-response part of the assessment.

Finally, tier 1 of the risk assessment is not a risk assessment at all. It is a hazard identification. I presume one would need to do a hazard identification as prelude to the assessment process. Overall, then, I recommend just two tiers for the entire assessment rather than this system of separate categories of tiers for separate parts of the assessment.

3. On Pages 5 and 6, I am generally supportive of the approach mentioned for proportional "roll down" of concentrations, so long as one can assume that control strategies really would affect all geographic areas equally (which I doubt, but the error introduced will not affect the fraction of population at or near a benchmark health effect level). But I am not sure about the utility of a "roll up" procedure based on the historical data, since it is not clear to me how one would determine which particular past historical data are most representative of what conditions will be like overall once a new regulation is in place. I'm not saying the idea is intrinsically wrong, only that I don't know how it would be executed.

4. On Page 7, the authors speak of the "relative degree of confidence". I have no idea what this means. In the same paragraph, they refer to a criterion (for moving to a more detailed and quantitative uncertainty analysis) if such an analysis adds "value". No coherent explanation of what "value" means in this instance is given, either here or later in the document. It might mean either that it better informs the decision or that it leads to different regulatory results. In the latter case, however, I don't see any discussion of how uncertainty relates to any kinds of decisions that might be made, and so it is not clear how one is to decide "value" in this utilitarian sense.

5. Beginning on Page 8, I began to have a problem with understanding the role of the tiers of the assessment. At first, I thought tier 1 might be a kind of screening assessment in which the assessor is asking: If I make several simplifying assumptions that tend to all overstate the risk, do I see any evidence of a significant risk? If yes, I will go to tier 2. If no, there is no need for me to proceed with any more detailed analysis.

But then the document describes the choice of moving to tier 2 as being related to the availability of data, and not to any specific results one sees from tier 1. So tier 1 does not seem to be getting used as a screening tool. I can't understand why one would even do tier 1 if the data are available for tier 2.

6. I am supportive of the use of PMR values to get at the short-term exposures in the geographic areas where only longer-term averages are available. I presume the assessors will develop CDFs for PMRs under different conditions (near sources, away from sources, etc) and apply the appropriate CDFs to non-monitored areas. The document at least hints at this, even if it is not expressed well. An example is on Page 10, where the bulleted list evidently applies to this issue, but the reader is not told why these four bulleted issues are being presented, or how their answers would affect the development of and application of CDFs.

7. On Page 13, the authors describe (at the bottom) an issue of 10 or 15 minute averages. It is not clear if this is to be a rolling average from the 5 minute predictions, or whether new PMR CDFs would be developed starting with the original monitoring data.

8. On Page 14, the authors appear (at the bottom) to be saying that measurement error is small compared to other sources of uncertainty. I would in general agree, but there will need to be some evidence of this before this source of uncertainty is ignored.

9. In several places, including on page 15, the authors mention a kind of sensitivity analysis to be performed, and then state that they will determine whether a given parameter or term does or does not contribute to uncertainty. All parameters and terms and models contribute to uncertainty. I assume they mean something like "contribute significantly".

10. I was not sure how results less than the MDL or MQL will be factored into the analysis of PMR distributions. Perhaps only results above the MDL or MQL will be used?

11. I am assuming that uncertainty factors will not be incorporated into any Health Benchmarks used. If they are, then this will need to be reflected in the uncertainty analyses.

12. There is a very general issue I want to raise concerning the incorporation of activity levels in the assessment. To the extent the clinical data are used, this makes sense, since the effects at a given concentration are tied to activity level. So it will be necessary to estimate the activity level of an individual in the exposed population to determine which clinical exposure-response curve to use. But for the epidemiological results, variations in activity level are already hidden inside the slope factors. In fact, the slope factors at low exposures probably are driven by the fraction of people who are both sensitive and exercising in a population at the time the study was done. So it might not be appropriate to do a detailed exposure assessment, complete with inter-subject variability of exercise patterns, and then apply the slope factor or other risk summary from an epidemiological study to all exposed individuals regardless of activity level. Having said this, however, I am not sure the authors intend to do this anyway, since I cannot understand from the document HOW they intend to use the epidemiological results.

13. There are two ways to use the monitoring results for air measurements in conjunction with dispersion models. One is to calibrate the models to the data. The other is to use the data in a model-to-monitor comparison for purposes of uncertainty analyses. The authors appear to be leaning towards the latter, but this isn't stated clearly. In any event, I would prefer the former.

14. On Page 19, I assume the modeling will allow for overlap of plumes from multiple sources in a geographic area. This isn't stated.

15. I am generally supportive of the use of APEX. The one caveat I would apply here is that this may be more detail than is justified by the concentration-response results. And it will be difficult to defend the idea that any resulting PDFs of exposure reflect actual exposures on time periods as short as an hour or less. This is an area of assessment in which the uncertainties are very large due to the extreme variation of an individual's activities during a day. It will be important to

present the assessment as a scenario analysis of representative exposures in a hypothetical (but reasonable) population, and not as an accurate representation of actual exposures to individuals.

16. On Page 24, the authors describe the use of a national average for asthma prevalence rates. But if this were valid, it would imply that these rates don't depend on geographic location, which would in turn imply that they don't depend on levels of exposure to air pollutants, which seems to go in the face of the basis for many of the NAAQS standards. I am not saying this is a bad approximation, or even the best that can be done, but it does lead to a logical inconsistency.

17. On page 24, the authors use a phrase that appears often in the document: "…assessment would take into account…". I agree with the sentiment, but no guidance is given as to HOW or IN WHAT SENSE something will be taken into account.

18. On Page 25, the authors mention sensitivity analysis. I support the performance of such an analysis, but a decision must be made as to whether it will be a local SA (adjusting one parameter at a time) or a global SA (adjusting multiple parameters and looking at contribution to variance).

19. On Page 25, the issue is again raised of comparing model results to monitors. The problem here (which also appeared in NATA) is that a model may get a peak value correct but have it shifted slightly in space. So if one simply compares model results at a point against monitor results at the same point, an overstatement is obtained of the uncertainty in exposures to a population.

20. On Page 31, the authors use the phrases "source-oriented focus" and "urban-area oriented focus". I am not sure what these mean or why they are needed, unless the decision is whether to use a representative set of locations based on source type or a set based on general urban characteristics.

21. Also on Page 31, in the last paragraph, my lack of clarity as to what is being done with the epidemiological results makes it impossible for me to understand this paragraph. Again, the treatment of the clinical results is clear in the document (once one gets to Section 4, at least), but not the treatment of the epidemiological results.

22. On Page 33 and at several other points, the authors point to the need for baseline incidence. This is only true if a relative risk, rather than absolute risk, model is used. The epidemiological papers certainly report relative risk summaries, but they also report the primary data from which absolute risk values can be calculated. Of course, one does not want to calculate absolute risk values if the biological processes are truly more consistent with relative risk (where the excess incidence is itself a function of the background incidence).

Comments from Dr. Schlesinger

p. 24. 3.3.4. First bullet should read “Healthy Children.”

p. 27. 1st paragraph. Health endpoints are not causal to ambient SO2. Rather, ambient SO2 is causal to health endpoints.

p. 28, 2nd paragraph. What is the criterion for judging whether or not a health effect that is considered to be of public health concern will not be appropriate for inclusion in quantitative assessment?

p. 30. Should elderly be included in bulleted list?

p. 35, 2nd paragraph. In the last sentence, it is noted that risk estimates may sometimes be developed using two different models. What will be the criteria for determining which model will provide some basis for evaluating the ultimate NAAQS?

Comments from Dr. Seigneur

The two-tier approach for exposure assessment and the three-tier approach for risk assessment appear to be logical ways to proceed. The various steps of each approach are described with sufficient detail for the reader to understand the technical approach and the sources of the data to be used. The use of AERMOD for the Tier 2 exposure assessment is appropriate.

Emissions vs. concentrations:

The discussion of the Tier 1 exposure assessment (Section 3.2.1, p. 8) focuses on the largest emitters. Clearly, the analysis must address the largest SO_2 emitters, but one must keep in mind that a smaller emitter with a short stack may have a greater impact in terms of SO_2 ground-level concentrations than a large emitter with a tall stack. The atmospheric dispersion aspect of the exposure assessment will be addressed explicitly in the Tier 2 exposure assessment but the potential limitations of the Tier 1 assessment must be clearly stated when the results are reported.

Figure 2 presents emissions by source categories. The year of this emission inventory should be stated because some source categories (e.g., coal-fired power plants) are being controlled and the emissions of those source categories will decrease. Also, one must note that some source categories, which may appear small in a nationwide inventory, may be quite relevant in an exposure assessment because they are concentrated in a few geographical areas (e.g., ocean-going ships in ports).

Areas of interest for exposure assessment: In the first paragraph on page 10, it is stated that cities in California report the lowest mean concentrations and that cities in the Northeast report the highest. This result is consistent with SO_2 emissions from coal-fired power plants being historically greater in the Northeast than in the West. However, this result depends strongly on the locations of the SO_2 monitors. One may assume that the monitoring network was designed to track the impact of SO_2 emissions from coal-fired power plant emissions. As this source category is being controlled, other source categories may become of concern and the existing monitoring network may not characterize their impacts properly. For example, SO_2 emissions from ocean-going ships could lead to significant SO_2 concentrations in ports (they typically burn 1.5% sulfur content fuel during transit within sulfur emission control areas, SECAs) but monitors may not be located strategically in those areas. This point should be kept in mind for the Tier 1 exposure assessment and the potential impacts of ship emissions in areas where those emissions are concentrated (ports and channels) should be explicitly addressed in the Tier 2 assessment. For example, some port areas could be selected for detailed concentration modeling (e.g., Houston, Los Angeles) under the Tier 2 assessment.

Comments from Dr. Frank Speizer

Air Quality Considerations

1. Use of proportional approach to adjust simulated potential alternative SO2 standards.

This seems reasonable but there is an issue (see next comment)

2. Use of historic air quality data pre 2000 vs. proportional approach. .

It is not an unreasonable use of historical data. However, with sites quoted in which there are extremes of Policy Relevant Background that can be less than 1% (Ohio Valley) vs. >70% (Volcanic region) it is not clear that there is a simple alternative approach. Additional regional consideration will need to be played out.

Exposure Analysis

1. Tier I approach

. End of 2nd paragraph on page 7: It is not clear what is meant by: ".a quantitative assessment of uncertainty …for selected components of the assessment".

Section 3.2.1 Approach seems fine but there ought to be a model approach that would allow for some validation that what is being done to make the proportion estimates of 5 minute maximums of <0.5ppm (or <0.4 or any other max). Given the limited number of co-located 5 minute and 1 hr samplers some "jack-knife", or other multiple statistical estimate of degree of concurrence ought to be possible.

Page 8-9. Given that electric generation ~75-85% of SO2 emissions, and virtually all of this is point source generated, there must be enough modeling data to allow for some generally hourly predictive modeling by distance from source (e.g. <20Km, 20-40Km, >40Km). Given such data and the possibility of co-location with some continuous monitors, I do not think Staff should block these data into yearly assessments but should use the co-located data across all years for the model development and then use trends over time to assess frequency of 5 minute max's over given levels (*as is* and other alternatives). I would be most disappointed if Staff concluded that they could not get past Tier I.

Page 14, end of first full paragraph. It appears that Staff has not yet decided it plans to use 5 minute max's of 0.5ppm or 0.6ppm or the criteria for selecting even lower levels. Factors influencing exposure levels needs some more discussion on how the choice will be made. One issue that should come up later in the health risk discussion relates to the potential sizes of subgroups that will be either susceptible or vulnerable. Back on page 3 Staff estimates that 0.7-1.8% of total asthmatic population could be exposed to outdoor SO2 > 0.5ppm for >5 minutes while exercising. That translates to a very big number given 10 million US asthmatics!

Page 16, first bullet, sentence beginning 5 lines from end: Not clear what this is saying. Aren't ambient and outdoor concentrations the same?

Criteria for assessing the uncertainty seem appropriate. I would add an additional bullet that states that some kind of validation will be implemented and that some criteria of validation are used to accept or reject the modeling be included before moving to Tier II (I am assuming that such modeling will be acceptable and that Tier II will be performed.

3. Tier II exposure assessment

a. Modeling SO2 by proximity to sources. See above. This is a very reasonable approach given the vast majority of source emissions are concentrated in stationary sources.

b. Binning. The bins seem appropriate but it would be worth obtaining population exposure estimates downwind from these bins in perhaps 3 tiers of <20, 20-40, >40Km of distance.

c. Using peak-to-mean ratio cumulative density functions. This seems reasonable but may not go far enough. Given that a single peak over 0.5ppm in an hour in the past predicted at least another peak in the same hour over 70% of the time and currently predicts about 35% of the time, does more thought needs to be given to modeling multiple exposures? Alternatively, does it suggest that additional modeling needs to be done to estimate how much the hourly average needs to be lowered to have 1 or less peaks above a certain value?

Top of page 23, sentence beginning end of line 5.and discussion in next paragraph. Surely time-location activity patterns must be almost random within waking hours. The estimates of overlap of activities within any given hours must relate to: 1) being outdoors (otherwise getting half the dose); 2)Exercise; 3)Frequency of 5 min averages over 0.5ppm; 4) somewhere between 35 and 70% of the 12 5 minute averages being over 0.5ppm; 5) other factors.

Section 3.3.4, page 24. Population groups of interest. Although this is an improvement over what was offered for NO2, if possible I think it would be useful to consider children broken down further. I think it would be better to consider birth- preschool (near home); 4 or 5 to 9 (local community); and 10-18 (active outdoor physical activity). I recognize that the data may not exist but at least the breakdowns for exposure might be considered. The other groupings seem appropriate, except might want to consider those adults carrying a chronic respiratory disease and or CVD diagnosis as a separate (potentially more susceptible) group.

4. General approach to uncertainty.

Page 25, last paragraph before section 3.4. This paragraph discusses estimating model uncertainty and suggests relying on “informed judgment” It is not clear whose judgment is being relied upon. Is this related to the desire mentioned elsewhere to conduct “expert opinion assessments”?

Health Risk Assessment

1. General Structure seems appropriate except for what I would consider one major omission. Throughout this discussion little assessment is given to the idea that there are within almost all population groups subsets of particularly sensitive people. I am not talking about the identified susceptible or vulnerable. Going back decades exposure studies in otherwise normal people always identified individuals who were more sensitive to SO2 than the group being studied. (Generally this amounted to 10-20% of the population being examined). The same seems to be true in population studies and as one moves to exercising adults, asthmatics and exercising asthmatics the percent susceptible or vulnerable increases. Often in the general population studies these particularly susceptible people are overwhelmed by the larger groups and the relationships for the group are considered null. They obviously are not null and in conducting a risk assessment that is supposed to take into account a margin of safety for the particularly susceptible I fear the effects in these groups are being downplayed.

Finally, once again Staff has included on page 28, first paragraph in there decision matrix of whether to conduct a Tier III risk assessment an unacceptable criteria of time and resources to complete the task. If a Tier III assessment is warranted this cannot be a criteria for not doing it!

2. Tier I health risk assessment

Questions a-c. Agree with planned assessment as interpreted from the initial draft of the ISA that the primary focus should be on the respiratory outcomes as described. However, in the ISA there were rather convincing evidence that for short term exposures in adults with pre-existing disease that cardiovascular short term effects were present. This will need to be revisited after further discussion of the ISA.

3. Tier II risk assessment

a. Agree with potential respiratory health risks for the controlled exposure studies. However, these are mostly designed to assess and understand potential mechanisms for the risk observed in free-living populations. The plan as outlined for a two Tier effort, like for the exposure assessment, seems somewhat arbitrary as to whether it is called a two tier effort or a logical progression in gathering the data necessary (and I believe from the draft ISA) available to do all that is proposed. The short term exposure assessment is well documented to move forward, particularly for the respiratory outcomes described. With regard to the long term assessment particularly for hospitalizations and mortality by sub-regions, this may have to await the assessment of the draft ISA.

b. Selection of benchmarks of 0.5-0.6ppm. I am concerned that these may be too high, particularly because as indicated at a minimum these values also predict a second or greater number of times during the same hour at least 35% of the time. For asthmatics exercising outdoors this is an unacceptable exposure. Also there are studies in asthmatics that show responsiveness below 0.5ppm.

c. Focus on areas around major sources. Yes, but will need population estimates by distance from point sources.

d. Yes, however, Staff needs to quantify those places in urban areas that are within some defined downwind distances (<20Km) to significant point sources.

e. Once again the ISA should have summary tables that provide data on exposure parameters for the studies to be used in assessing averaging times, location and season in addition to co-pollutants. Working with staff in producing these table will provide data needed here. I do not suggest that those tasked with doing this risk assessment go off independently to construct such tables, since this might become an excuse for not doing the assessment that needs to be done (because of lack of time or resources).

4. Tier III risk assessment

a. I do not agree that there is insufficient information to develop a credible exposure response relationship from controlled human exposure evidence. SO2 is one of the most and best studied pollutants in terms of controlled human exposure. The data base is quite rich and with careful assessment has been used to assess a variety of dose-response relationships and has identified susceptible and vulnerable subgroups. I believe this will come out in our assessment of the ISA. At least I would reserve judgment on this point until after review of the ISA.

b. Here again I believe the success of modeling of exposure from the 5 min max from the l hour will dictate whether it would be useful to consider using the epi data for a Tier III assessment of short term effects. For longer term effects, consideration of identifying

susceptible subgroups (e.g. adults with preexisting disease, as a result of long term residence in highly polluted regions or with repeated 24 hour exposures) will need to be assessed further.

Comments from Dr. Wyzga

Overall comment: It is difficult to describe and evaluate the plan generically. At times I personally cannot understand the methods as described. A fuller explanation with examples would enable me to judge them better. Given the deadlines faced, however, it would be impractical to redescribe the methods with more detail. The best approach would be to apply them and alter them, if necessary, given various reviews of their implementation.

Health Risk Assessment

Question 1: It seems to me, from reading the ISA and from following the literature, the risk assessment targets the correct health responses, those of asthmatics with relatively short exposures. Given this, I'm not convinced that there will much value in placing much additional effort and resources in a Tier I Health Effects Evaluation.

Question 2 a: See above; I would not place much additional effort on this qualitative, but proceed to Tiers II and III.

2b: I agree; since these studies are performed with controlled exposures, there is less concern about confounders. In addition, these studies find responses after exposures as short as 5 minutes; I am unaware of any epidemiological study that has or can adequately address such exposures.

2c: I agree – exercising asthmatics appear to be especially vulnerable.

Question 3 a. I believe so.

3b. this seems reasonable

3c. yes

3d. I worry about the short-term SO2 issue and whether urban monitoring data are really characteristic of exposures. The approach is clearly more reasonable for the 24-hour exposures. I have no easy solution to offer for the one-hour problem. Near-source exposures may be more important for one-hour exposures and should be considered.

3e. I'd be happy to share any data collected by EPRI and its Contractors. I think the Agency has to be aggressive in seeking such data. The States may have more detailed data than has been reported.

Question 4a. Risk assessments can be undertaken for specific sources. They could be undertaken as illustrative. I would note that I was a co-author of one such risk assessment several years ago. See: P. C. Freudenthal, H. D. Roth, T. Hammerstrom, and C. Lichtenstein, "Health Risks of Short-Term SO_2 Exposure to Exercising Asthmatics", JAPCA, Journal of the Air and Waste Management Association, 39(6), June 1989.

4b. I would urge the panelists to consider breaking up any study area into small geographic units. I worry that effects may be seen in an area because several individuals were exposed to a plume with much higher concentration levels than measured by the monitor. The monitor values may reflect the existence of a plume, but at levels much lower than the plume; hence health effects would be detected, but in reality these effects are ca7used by exposures to the

higher plume values. This is not an easy issue, and there may be no easy solutions; hence whatever approach is taken, it is important to state caveats and limitations of the analysis.

Question 5: To the extent possible I would urge that the consideration of uncertainties be embedded into the risk analysis rather than undertaking a baseline analysis and several sensitivity analyses. The result should be some distribution of estimates.

If I recall correctly, chamber study effects were seen in exercising asthmatics who were medication-free. This factor could be considered in the overall risk assessment because medication is taken to protect from responses to lots of environmental agents in addition to SO2 and air pollution.

Comments from Dr. Larson

Air Quality Considerations:
Based on the low estimated contribution of policy-relevant background SO_2 to overall ambient SO_2 levels, staff is considering a proportional (i.e., linear) approach to adjusting air quality to simulate just meeting potential alternative SO_2 standards that are below recent air quality concentrations. Do the Panel members have comments on adopting a proportional approach to simulate just meeting more stringent alternative air quality standards?

This is a reasonable approach, given the very low policy-relevant background concentrations in most areas of the U.S. Although there are a few areas affected by natural sources (some locations in the Northwestern U.S.), the current levels are very low.

Recognizing that current ambient air quality concentrations are lower than the current standards, the draft Health Assessment Plan discusses two alternative approaches to simulating ambient SO_2 levels associated with just meeting the current SO_2 standards: use of historical air quality data (e.g., possibly pre-2000) when ambient levels were at or above the current standards, or use of a proportional (i.e., linear) approach to adjust SO_2 levels upward. Do the Panel members have advice or comments on these two alternative approaches to simulating air quality just meeting the current SO_2 standards?

The historical data has the theoretical advantage of including any non-linearities in the model. It is not clear that this is an actual advantage (one major non-linear effect would be due to a large policy-relevant background value, which does not exist). Another drawback is that the historical approach would capture a different mix of sources than currently exists. Therefore the relative shape of the distribution of short term values could be different than it is now.

Exposure Analysis:
1. In considering the exposure analysis broadly:
a. Do Panel members have any comments on the general structure and overall two-tier approach that staff plans to use for the exposure analysis? Are the criteria that staff plans to use for deciding whether to conduct a Tier II analysis clear and appropriate?

The criteria seem reasonable and practical.

b. Have the most important factors influencing exposure to SO_2 been clearly accounted for and described?

Given the dynamic complexities of plume dispersion, it is reasonable to capture some of this with a deterministic model such as AERMOD. The daily changes in mixing depth strongly influence the downwind impacts from a given elevated source.

c. The draft plan describes the basis for and selection of population groups of interest

(i.e., children, asthmatics (children and adults), and the elderly) for which SO_2 exposure estimates are to be developed. Do Panel members generally agree with the groups of interest identified in the draft plan?

This seems to be consistent with the ISA.

2. In considering the Tier I exposure assessment:
a. Do Panel members agree that a statistical model using available ambient 5-minute monitoring data is appropriate for estimating expected exceedances of very short-term (5-minute) potential health effect benchmarks?

There is relatively limited data on the 5 minute values. However, there is no reason to distrust what little data there is.

b. Do Panel members agree with the approach of applying a statistical model to estimate 5-minute concentration exceedances at monitoring locations where only 1-hour monitoring was performed for evaluating the extent of 5-minute peaks associated with meeting alternative standards with longer averaging times?

If the model is based on historical data or on locations that are qualitatively different than the locations of interest (e.g. point source impacted vs general urban), then care should be taken to make sure that the potential source impacts are appropriately considered. Another issue is the source geometry-tall stacks vs near ground level releases. These two geometries can, in principle, lead to different peak to mean values, especially during convective daytime conditions (c.f. Van Dop et al *Boundary Layer Meteorology* 116, 1-35, 2005; also Luhar et al. *Atmos. Environ*. 34, 3599-3616, 2000; Franzese and Borgas *J. Appl. Met.*, 41, 1101-1111, 2002).

3. In considering a potential Tier II exposure assessment:
a. Do Panel members agree with the combined emissions/dispersion modeling approach to estimate short-term (hourly) SO_2 concentrations in close proximity to SO_2 emission sources?

The approach described in equations 1 and 2 is an attempt to allocate the distribution of 5-minute SO2 values within in a given hour. Some thought might be given to the use of the standard deviation of wind direction for each hour, if such data are available. This is an indirect measure of the breadth of this distribution and is invoked in odor models. The effects of narrow plumes that have recently passed over water (relatively smooth surface) would narrow the 5-minute distribution. These effects are not predicted by AERMOD (as far as I know), but could be crudely captured by geographical variables.

b. Do Panel members have comments or advice regarding the described binning of sources and development of prototype stacks/facilities?

It might be useful to look at the measured peak to mean ratios as a function of wind direction. This could be done using a conditional probablility function, i.e., looking at the probabilities of occurance from a given direction when the ratio exceeds a certain value. If the high peak values are due to major point sources, this approach could "point" to the source's location. In turn, this might allow one to segregate point source-influenced measurements from those influenced by more widely distributed sources. Also, as stated earlier, the release height is an important determinant of peak to mean ratios.

c. Do Panel members agree with the approach using peak-to-mean ratio cumulative density functions (PMR CDFs) to estimate very short-term peak concentrations from the 1-hour modeled concentrations?

Seems reasonable, especially allowing these values to vary with 1-hr levels.

d. Do Panel members generally agree that the approach described using APEX is reasonable and appropriate to estimate the occurrence of very short-term (5 minute) SO_2 peak exposures?

In general, it is a reasonable approach. I cannot tell how sensitive it is to the assumptions of the actual 5-minute distributions for any given hour. It would seem important. Perhaps this can be included as part of the uncertainty analysis.

e. Do Panel members have any comments or advice regarding the general approach to addressing uncertainty and variability in each Tier of the exposure assessment as described in the draft plan?

Given all the uncertainties and lack of locations with 5-minute data, the approach of discussing the potential uncertainties is reasonable.

General Comment on Risk Assessment:

If possible, it would be good to try to include the fact that in most industrial parts of town, the incomes are lower and the access to managed care for asthma is poor. Yet these are the same locations that experience the highest peak concentrations from industrial emissions.

Comments from Dr. Thurston

This presented plan appears to sufficiently address the needs for the scope and approaches for, and highlights key issues in, the estimation of population exposures and health risks posed by SOx under: 1) existing air quality levels; 2) upon just meeting the current SO_2 primary NAAQS, and; 3) upon just meeting potential alternative standards under consideration by the Administration.

I do have several concerns, however. On page 2, the report limits itself to the effects of gaseous SO_2 alone, dismissing any role by particulate sulfates, or sulfur dioxide's interactions with PM in general, in assessing the possible health impacts of sulfur oxides. This ad hoc decision seems too yield too narrow a definition, and may cause an underestimation of the risks of sulfur oxides, as well as in the benefits achievable via sulfur oxide emissions reductions.

Another concern that I have is in regard to the reliance on multiple pollutant models for risk assessment, as proposed in the middle of page 35. This, despite the fact that the report states in the same paragraph: "When collinearity exists, inclusion of multiple pollutants in models often produces unstable and statistically insignificant effect estimates for both SO_2 and the co-pollutants.". Thus, a reliance on published single pollutant model coefficients would seem far preferable.

With regard to the Health Risk Assessment Charge Question 3.b. (Do Panel members generally agree with the tentatively identified potential health effect benchmark of 0.5 to 0.6 ppm for exercising asthmatics following 5-10 minutes SO_2 exposure?), I would say that I feel that this is too high, based upon my reading of the ISA draft. This ignores evidence and biological plausibility regarding the lowering of threshold by the co-presence of PM, which is always the case in the environment, and would also provide no margin of safety vs. the clinical study results. I should think a benchmark closer to 200 ppb would be more appropriate and of more interest to CASAC.

Finally, with regard to Health Risk Assessment Charge Question 1, I must object to the inclusion of time and resources as a criteria for determining whether to do a Tier II analysis (pg. 36, last two lines). This Tier III analysis needs to be done, and has this been known about by the EPA for years, so this should not appear as a criteria. Drop this last bullet from the list.

Comments from Dr. Kenski

Air quality considerations:

Q.1: With respect to charge question 1, the proportional approach for simulating concentrations below recent data to examine scenarios that just meet alternative standards is acceptable.

Q2: I have a slight preference for using historical data to examine scenarios with concentrations above current standards, because it may be that the historical data have subtle differences in distribution that would not be captured by the proportional 'roll-up' approach. Alternatively, an analysis of distributional differences could be performed to demonstrate that distributions from higher-concentration historical conditions do not differ appreciably from current lower concentrations.

Exposure analysis:

Q1. The general structure and approach are logical. It's actually not clear what the criteria for deciding to conduct a Tier II exposure analysis are, versus a Tier 1 assessment. The document is quite clear about how and what will be done, but it seems to implicitly assume that both will be performed (and there doesn't seem to be any reason not to perform the exposure assessment through Tier II). The factors in Sec. 3.4 are a little vague. If the ambient air characterization leads us to believe that no current ambient concentrations are above any potential alternative standard, we're done? No further exposure or risk assessment is necessary? The important factors influencing exposure and populations of interest have been accounted for.

Q2. I liked the proposed model for estimating peaks at monitors with 1-hour data. Until it's actually tested with some of the sites where 5 min data are available, it's not possible to give it an unqualified approval, but it seems eminently reasonable for generating the needed data. Of course the exposure assessment will need to document the performance of this model and document its contribution to the overall uncertainty assessment.

Q3. The Tier II approach made a lot of sense. I wonder, however, if the choice of most recent 3 years of meteorology is necessarily best? Is there any evidence to indicate that years vary significantly in their potential to be more or less conducive to high SO2 concentrations, independent of changes in emissions? E.g., perhaps cooler summers have slower SO2->SO4 conversion and so SO2 concentrations are higher at near-source monitors? Maybe an examination of yearly CDFs or quantile-quantile plots would show year-to-year differences. If so, then perhaps an argument could be made for selecting years that are more likely to have higher SO2, to ensure that modeled concentrations would be conservative.

The binning approach is a reasonable one; the development of bins is an interesting problem in itself. The inclusion of terrain as a variable is important; it's not clear whether this is definitely going to be incorporated or just examined as a possible option.

Q4: The discussion of uncertainty was helpful.

Health Risk Assessment:

Q1: The approach to the health risk assessment was clearly laid out and reasonable. The criteria for conducting a Tier III assessment were more clear than those for the exposure assessment section (especially the 1st paragraph on p. 28). Also, the 2nd paragraph on p. 28 was a particularly nice description of the goals of this process.

Q2: Based on the data presented in the ISA, I agree with the staff's assessment of the health endpoints and susceptible populations of most interest.

Q3: I have no expertise in health risk assessment, so I can only answer these questions based on what was presented in the ISA. That said, I agree with the staff's choices with respect to health benchmarks and the focus on exercising asthmatics. Certainly the decision to focus on areas around major sources and on urban areas is appropriate.

Q4 & Q5: No additional comments.

Sec. 3.2.1, 1st paragraph: This description of monitoring could use some additional clarification. Do the 94 monitors that report 5 minute maxes report one maximum 5 minute concentration per hour (or day?), or 12 5-minute values per hour (is this what is meant by 'containing continuous monitoring'? Even if AQS only contains one 5-minute max per hour, the states or local organizations that collected the 5 minute data may have archived measurements for the other 11 5-minute intervals. Please be sure to check with them, since the number of monitors is limited, to see what additional measurement data might be available.

Sec. 3.2.1, 2nd paragraph: Electric generating units are the largest source of SO2 nationally, but on a local scale many other sources are significant – industrial coal use, refineries, coking, metal processing, paper mills, and shipping (bunker fuel use). The proximity of these types of sources to the monitors will need to be considered in the analyses proposed in Secs. 3.2.1.1 and 3.2.1.2, not just EGUs. The last sentence of this paragraph makes it sound like proximity to these other sources may or may not be accounted for in the data analyses. (this seems to be addressed adequately in later sections of the report, just not right here)

Sec. 3.2.1.4, p. 14, 2nd paragraph: The application for this analysis of population density isn't clear. Will these estimates of susceptible populations be generated just for the vicinity around each ambient monitor or scaled up for the nation?

Comments from Dr. Gordon

The Plan is well conceived and written and the tiered approach is appropriate for the task. Because of a lack of expertise on exposure assessment and modeling, I will comment only on the health portion of the risk assessment. The conclusion that adverse respiratory effects are the strongest health findings appears to be valid and clearly substantiated by the ISA. The advancement of the respiratory hospital admissions and ER visits to a Tier III analysis is needed and verified. It is puzzling, however, why the Assessment Plan indicates that while there is clear evidence of bronchoconstriction in asthmatics after short term exposure to 0.5 to 0.6 ppm sulfur dioxide, it has been decided not to do a Tier III evaluation on this health effect. The dose response for acute bronchoconstriction has been know for nearly 2 decades and a Tier III evaluation of this health effect is warranted. As stated for Tier II evaluations, the approach may be different for epidemiology and controlled human exposure studies, but the quantitation of acute data from controlled human studies is feasible. If EPA feels that this quantitative assessment is not possible, then additional justification is required.

The Plan states that a Tier III risk assessment depends on a number of factors including "whether or not there is adequate time and resources". Given the enormous effort and resources (time and money) used to put together the ISA and the Assessment Plan (funded research, scientific review of grants and publications, EPA scientists writing the ISA and the Plan, CASAC panel members' review process, etc.), it is unclear why resources may not be available to accomplish this last and critical step in a timely fashion.

Comments from Dr. Hattis

1. Based on the low estimated contribution of policy-relevant background, and ambient SO2 levels, staff is considering a proportional (i.e., linear) approach to adjusting air quality to simulate just meeting potential alternative SO2 standards that are below recent air quality concentrations. Do the Panel members have comments on adopting a proportional approach to simulate just meeting more stringent alternative air quality standards?

> This seems generally reasonable to me.

2. Recognizing that current ambient air quality concentrations are lower than the current standards, the draft Health Assessment Plan discusses two alternative approaches to simulating ambient SO2 levels associated with just meeting the current SO2 standards: use of historical air quality data (e.g., possibly pre-2000) when ambient levels were at or above the current standards, or use of a proportional (i.e., linear) approach to adjust SO2 levels upward. Do the Panel members have advice or comments on these two alternative approaches to simulating air quality just meeting the current SO2 standards?

> To the extent possible, the goal should be to represent a realistic future scenario—one that might actually occur. One such scenario would be a generalized increase in present emissions resulting from increased SO2-emitting economic activities of all kinds. It seems likely to me that this would approximately correspond to the proportional (linear) approach rather than the historical reconstruction.

Exposure Analysis:

1. In considering the exposure analysis broadly:
a. Do Panel members have any comments on the general structure and overall two-tier approach that staff plans to use for the exposure analysis? Are the criteria that staff plans to use for deciding whether to conduct a Tier II analysis clear and appropriate?
b. Have the most important factors influencing exposure to SO2 been clearly accounted for and described?
c. The draft plan describes the basis for and selection of population groups of interest (i.e., children, asthmatics (children and adults), and the elderly) for which SO2 exposure estimates are to be developed. Do Panel members generally agree with the groups of interest identified in the draft plan?

> Yes.

2. In considering the Tier I exposure assessment:

a. Do Panel members agree that a statistical model using available ambient 5-minute monitoring data is appropriate for estimating expected exceedances of very short-term (5-minute) potential health effect benchmarks?

b. Do Panel members agree with the approach of applying a statistical model to estimate 5-minute concentration exceedances at monitoring locations where only 1-hour monitoring was performed for evaluating the extent of 5-minute peaks associated with meeting alternative standards with longer averaging times?

Generally the idea of modeling the 5 minute peaks with the aid of empirical data and a statistical model is a good one. I have not grasped the exact statistical model to be used sufficiently, however, to be sure that it fully realizes the opportunities presented by the available data and takes precautions to correct for the artifactual spreading of the data from measurement error. On the latter issue, it is somewhat troubling to see the discussion to the effect that only "valid" measurements will be used. Fine, and impossible 5 minute/hourly PMR values less than 1 or greater than 12 will be excluded. But this does not mean that the effects of residual measurement error in spreading out both the 5 minute and 1 hour average observations have been excluded. Any set of empirical observations has measurement error. In general the observed lognormal variance will be the sum of the real lognormal variance of real SO2 levels and some lognormal variance attributable to measurement errors. However only real variation affects real people's exposures and risks. Thus to get an estimate of the true frequency of high values of the exposure distributions (and the corresponding ratios of 5 minute/1 hour levels) it is important to estimate the measurement error variance (likely different for the shorter vs longer averaging times) and subtract that from the variance of the crude observations.

3. In considering a potential Tier II exposure assessment:
a. Do Panel members agree with the combined emissions/dispersion modeling approach to estimate short-term (hourly) SO2 concentrations in close proximity to SO2 emission sources?

Yes. I do, however, think that to the extent possible some effort should go into comparing observed and dispersion model predicted distributions of hourly SO2 levels at monitors near specific sources. Based on the results of this comparison, the distribution of hourly SO2 levels for unmonitored sites may be adjusted for better accuracy.

b. Do Panel members have comments or advice regarding the described binning of sources and development of prototype stacks/facilities?
c. Do Panel members agree with the approach using peak-to-mean ratio cumulative density functions (PMR CDFs) to estimate very short-term peak concentrations from the 1-hour modeled concentrations?
d. Do Panel members generally agree that the approach described using APEX is reasonable and appropriate to estimate the occurrence of very short-term (5 minute) SO2 peak exposures?

Yes, generally to b, c, and d, subject to my earlier comments about the need to separately remove the effects of measurement errors from the 5 minute and hourly data that give rise to the PMR CDFs.

4. Do Panel members have any comments or advice regarding the general approach to

addressing uncertainty and variability in each Tier of the exposure assessment as described in the draft plan?

The second paragraph on page 7 says in part, "At each tier of the exposure assessment, an evaluation of the uncertainties will be performed and the relative degree of confidence in the exposure estimates will be determined." "Determined" is a bit stronger word than I would like to use in general for an uncertainty analysis. Consider substituting the more modest terms, "estimated" for a quantitative analysis, or "assessed" for a more qualitative or semi-quantitative discussion.

Health Risk Assessment:

1. Do Panel members have any comments on the general structure and overall three-tier approach that staff plans to use for the risk assessment? Are the criteria that staff plans to use for deciding whether to conduct a Tier III risk assessment clear and appropriate?

I think so.

2. In considering the Tier I risk assessment:

a. Do Panel members agree with the approach of having a qualitative assessment of health endpoints to identify which are likely candidates for a more sophisticated and quantitative tier of assessment?

Yes.

b. Do Panel members agree with our initial observation that controlled human exposure studies demonstrate strong evidence for bronchoconstriction in exercising asthmatics following 5-10 minutes SO2 exposure?

Yes.

c. Do Panel members agree with staff's initial observation that the strongest epidemiologic evidence is for respiratory symptoms in asthmatic children and respiratory-related hospital admissions and respiratory-related emergency department visits in asthmatics and others with respiratory conditions?

Yes.

3. In considering the Tier II risk assessment:

a. In general, are staff plans to use potential health effect benchmarks to address respiratory effects demonstrated in exercising asthmatics in controlled human exposure studies clear and appropriate?

The proposal is clear. As with the NOx analysis I have reservations about the general plan to use "health effects benchmarks" and the incidence of exceedances as the main analytical approach. As I illustrated in my comments on the ISA (reproduced below), an approach that uses a crude log probit dose response function together with quantitative assessment of the full distribution of exposure concentrations is quite feasible.

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

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

For the analysis of human variability in Figure 2 below I have extracted the individual Horstman et al. data as best I could from the figure provided in the ISA and the accessible abstract (I could not easily obtain the original paper). Figure 2 is based on the conventional assumption for probit analysis that in the population of asthmatics studied there is a lognormal distribution of individual thresholds for the response (a doubling of airway resistance during exercise). In this case the intercept is an estimate of the log of the SO2 level needed to elicit the response in the median asthmatic (100.0189 = 1.044 ppm = 1044 ppb) and the slope is an estimate of the log of the geometric standard deviation [the Log(GSD) in our terminology] of individual response thresholds (100.374 = 2.37).

Figure 1

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

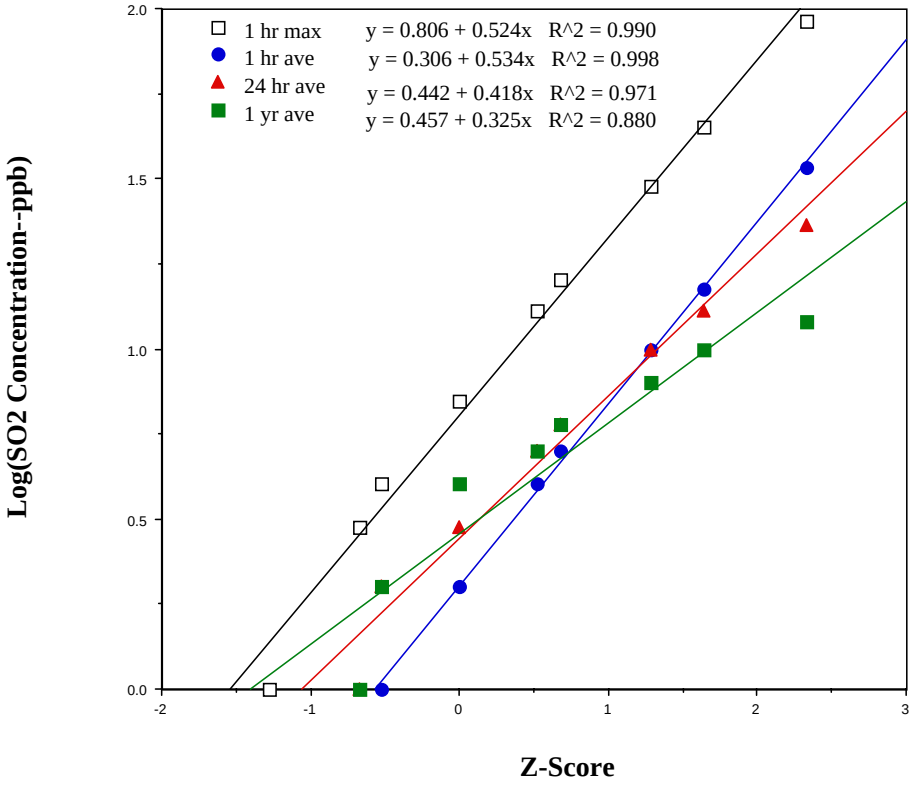
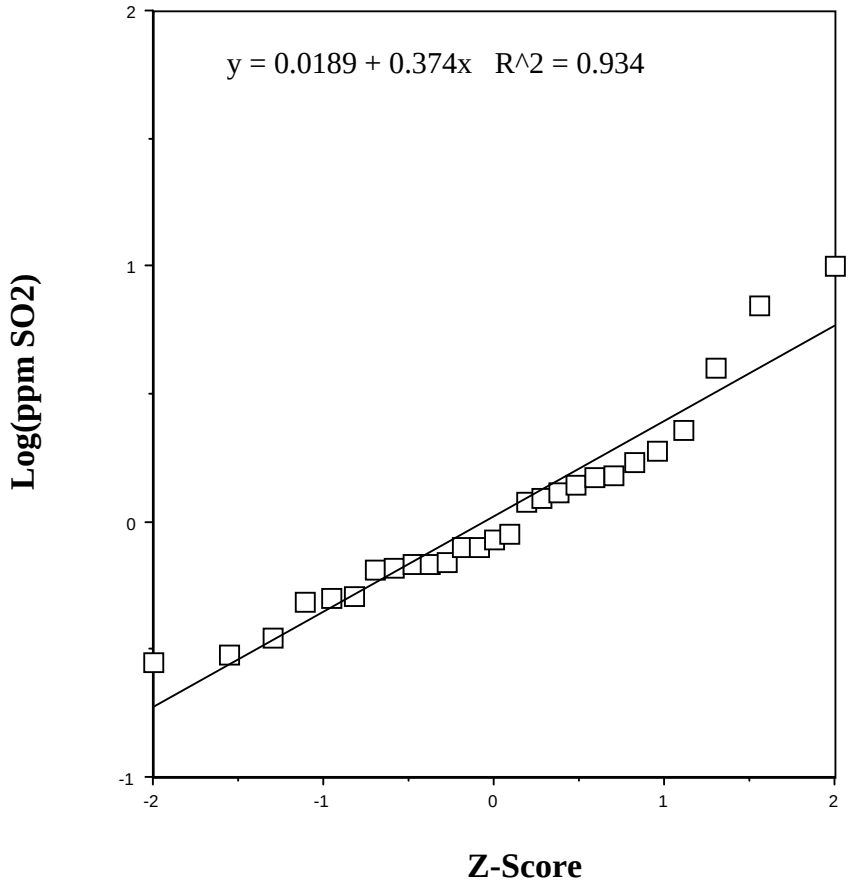


Figure 2

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



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

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

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

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

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

Source: Human interindividual variability database, updated as of 5/05, available "http://www2.clarku.edu/faculty/dhattis" discussed in
Hattis, D. "Distributional Analyses for Children's Inhalation Risk Assessments." Journal of Toxicology and Environmental Health, 71:1-9, 2008 in press.
Hattis, D. and Lynch, M. K. "Empirically Observed Distributions of Pharmacokinetic and Pharmacodynamic Variability in Humans-Implications for the Derivation of Single Point Component Uncertainty Factors Providing Equivalent Protection as Existing RfDs." In Toxicokinetics in Risk Assessment, J. C. Lipscomb and E. V. Ohanian, eds., Informa Healthcare USA, Inc., 2007, pp. 69-93.
Hattis, D., Baird, S., and Goble, R. "A Straw Man Proposal for a Quantitative Definition of the RfD," Drug and Chemical Toxicology, Vol. 25, pp. 403-436, (2002).

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

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

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

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

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

b. Do Panel members generally agree with the tentatively identified potential health effect benchmark of 0.5 to 0.6 ppm for exercising asthmatics following 5-10 minutes SO2 exposure?

No. Effects are clearly observed in some people well below this level, and the effect incidence for almost any level can be estimated (see responses above to the questions on the ISA), if one is willing to postulate an overall lognormal distribution of individual thresholds—which seems reasonably compatible with available data and applicable theory.

c. Do Panel members generally agree with the staff's approach of focusing on areas around major sources of SO2 with respect to concerns about 5-10 minute peak exposures related to the respiratory effects observed in controlled human exposure studies?

Yes.

d. Do Panel members generally agree with staff's approach of focusing on urban areas with respect to concerns about 1- and 24-hr and annual SO2 concentrations related to respiratory effects observed in epidemiologic studies?

Yes.

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

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

e. Do Panel members have any comments or advice with respect to staff's approach of gathering additional information to characterize the SO2 ambient air quality that existed at the time various key U.S. and Canadian studies addressing respiratory effects were conducted to see if the concentration-response relationships observed in these epidemiologic studies are related to particular SO2 levels and associated averaging times, geographic location and/or season, and the inclusion of various co-pollutants?

> I think this is an ambitious undertaking, but worth trying. The key issue of confounding might be addressed by trying to compare results of studies with more vs less vs different types of co-pollutant exposures, particularly organized by major sources of particulates in the areas studied by different authors.

4. In considering a potential Tier III risk assessment:

a. Do Panel members generally agree that there is insufficient information to develop credible exposure-response relationships for use in a quantitative risk assessment based on the controlled human exposure evidence?

> Not at all--the tables and figures I developed with only a couple of days of effort in the ISA response section above do exactly that for at least one type of response. A better job can be done with more efforts and a more sophisticated analysis, but surely some quantitative analysis of likely effect incidence is feasible.

b. Do Panel members have any comments or advice with respect to the general approach or specific factors to be considered in deciding whether or not to proceed to a Tier III quantitative risk assessment for the respiratory-related health endpoints based on epidemiologic evidence discussed in the draft plan?

> I think EPA should plan on doing a Tier III assessment for at least the simplest short term endpoints.

5. Do Panel members have any comments or advice with respect to the general approach to addressing uncertainty and variability in each Tier of the risk assessment as described in the draft plan?

> Just that it is important to treat at least variability in susceptibility quantitatively based on existing data in available clinical observation papers. Uncertainty analysis methods also deserve some quantitative attention.

Comments from Dr. Kinney

Air Quality Considerations:
1. I think the proportional approach for adjusting air quality is fine. However, I question the value and purpose of rolling up concentrations from ambient to the level of alternative standards. The only obvious reason to do that would be as part of a "benefits analysis" to demonstrate the health benefits of having ambient concentrations below the level of the standard. That's not the purpose of this exercise obviously, so why do it?
2. Note concern expressed above. However, if you must do this, I prefer the proportional adjustment method.

Specific Comments:
p. 5, para 3, last line: controlled exposure studies can provide useful exposure/response functions for use in risk assessment; this should be noted here.
p. 8, para 2, lines 1-2 and elsewhere: the term "surrogate exposures" is mentioned here and several other places, before any definition is provided. Need to add a couple of explanatory sentences early on to explain what is meant. It becomes clear later, but needs to do so earlier.

Exposure Analysis:
1.a. The general structure and process for the two-tiered approach is well justified and appropriate.
1.b. The most important factors influencing exposure to SO2 have been clearly accounted for and described.
1.c. The population groups of interest are appropriately chosen.

2.a,b. I think the statistical approach seems reasonable, although the description is somewhat unclear, as noted on in my comments on the draft document.

3.a-d. Modeling approach is reasonable. The binning of exposures sounds ok, but the devil will be in the details, and we'll need to see how well it works in practice. The PMR CDF approach is reasonable. I like the APEX modeling approach for getting at actual personal exposure distributions.

Specific Comments:
p. 10, para 1, 7th to 5th line from bottom: lack of correlation also likely reflects the high proportionate uncertainty for concentrations at or below the instrument LODs

p. 13, equations and last para: this material is a bit confusing. What is meant by "the appropriate function will be applied"? What is being estimated? Give an example calculation.

p. 15, 4th para, last sentence: This is hard to understand. Edit to clarify meaning. I had to read it several times.

4. Uncertainty approach makes sense in general.

Health Risk Assessment:

General comments:

Why would controlled exposure results not be useful for risk assessment?
Is risk assessment even warranted given the fact that concentrations are all below the standard?

2.a-c. Qualitative assessment is a good starting place. Agree with bronchoconstriction findings. With respect to epi, I don't find any of the epi data compelling and robust for SO2, although there are suggestions. The problem is that SO2 is too confounded by co-pollutants, and the levels of SO2 are far far below levels that have ever been observed to have relevant adverse effects in controlled studies.

3.a-e. All very well justified approaches.

4.a. No I do not agree with this. Unless I am mistaken, I don't think the document includes a rationale for this decision.
4b. More thought needs to go into deciding whether a tier III analysis would ever make sense based on the epi evidence alone.

Specific Comments:

p. 28, para 2, 4th numbered point: this one is a bit unclear; edit for clarity.

p. 29, para 3: although the ISA states that the SO2 effects were "generally" found to be robust, this contrasts with my interpretation of the results presented in the ISA. "sometimes" is a more accurate term to use regarding SO2 robustness. Also, I take issue with the ISA biological plausibility conclusion given the 2-3 order of magnitude higher concentrations at which the lab-based findings are seen.

p. 33, section 4.4, 1st para: justification for the statement that controlled exposure studies do not provide information to develop "credible exposure-response relationships" is nowhere to be found in the supporting materials up to this point, including the ISA. It is particularly surprising given the extensive attention devoted to 5 minute concentrations in the exposure work presented earlier. Why would one devote so much focus and effort on short term SO2 if there were insufficient information to develop credible exposure-response relationships ? This argument needs to be laid out carefully and convincingly. There may be a good argument, but it's not here.

p. 34, 2nd para from end: Given all the other uncertainties in this assessment, it is unreasonable to set the bar so high as to require same-location epi data before risk assessment can be conducted. It WOULD be preferable to have C/R data representative of the region (e.g., NE US), but I don't think it is essential to require even this in order to

do an assessment. The list of uncertainties presented here are all real, but no more problematic than those that appear in the exposure modeling for example.

p. 35, 2nd para: it should also be recognized that multi-city results may not be optimal for assessing effects in any one particular city and that uncertainties will be encountered if this is done.

p. 35, 3rd para, at the end: Another option would be to rely only on C/R functions from studies and models in which SO2 was included with co-pollutants AND where the SO2 effect was robust (i.e., the so2 effect did not change in going from single to multiple pollutant models).

p. 36, section on criteria for determining approach (numbering of section seems off), first bullet: need to be more clear about what is meant by "health effect benchmark levels associated with current ambient conditions." Are you suggesting that you'll use epi results to find thresholds?? You need to explain someplacehow such benchmarks would be determined from the epi data, since apparently it is only the epi data that would inform a tier III analysis.

p. 37, last line: Also should the proportion of the US population that is asthmatic, outdoors, and exercising while ambient concentrations reach 5 minute peaks of concern.

Comments from Dr. Russell

First, a number of issues that arose from my reading the ISA as to how the information from that document would be used down the road (e.g., in the Exposure and Risk assessments) were answered to a reasonable degree. Indeed, there seemed to have been additional forethought in the development of the Scope and Methods. Thus, at the first level, I am generally pleased.

In regards to the air quality and exposure assessment, the greatest concern I have continues on from the ISA, that being that our current monitoring system likely provides a relatively poor characterization of the levels of SO2 typically found in an urban area. The ranges of concentrations measured at any one monitor are very dependent on the geometry, particularly distance and direction to the monitor. If a monitor is close to, and from the direction of the dominant prevailing winds, of a major point source, it will have higher concentrations, and likely higher PMRs. On the other hand, a monitor further away will likely have lower PMRs. I wonder, would it not be easier to fit the data to a log-normal distribution and/or develop correlations between the 5 min. max (and possibly 2nd highest and third highest), along with the regression statistics, and use those to develop the relationships?

In regards to Tier I exposure assessment, you note that you will be looking at estimating peak 5 minute levels at monitors… why not populations?

This document seems to suggest that there are 5 minute monitors and separate 1-hr monitors. This strikes me as strange.

One of the questions that staff will have to examine is the probability of capturing the highest 5-minute average event in an area. One of the problems with this is that the observed PMRs are going to be highly dependent upon how close one is to the source, and the coexistence of many sources in a region. Thus, the PMR measured at a site may not be very representative of the area under consideration. This gets back to the issue of exposure error: one (or a few) monitors may be relatively poor measures of both average ambient air quality (much less personal exposure) and variability when the highest concentrations are due to plumes with a relatively small spatial footprint. It is interesting that a peak SO2 level of 600 ppb was measured at a monitor… this would indicate that the plume impacting that site had not dispersed much. Doing a mass balance would suggest that the plume cold not be very wide at that point.

The two statistical models (eq. 1 and eq. 2) need to be better explained, and the ramifications of the choice of model form spelled out. In eq (2), what are the individual Pi's, and how are they calculated? What do you mean that "m" is the number of peak concentrations? There can only be one. At this point, one wonders why not just impose a log-normal distribution? Indeed, I have a suspicion that the effort being described on top of page 20 can be done analytically if one assumes some sort of analytical form for the cdf of the PMRs and for the 1-hr concentrations. (I don't see any problem with such

an approximation, unless the data do not support such.) One could see the possibility of determining an empirical function where:

$$c(t_i) = f(\bar{c}, i)$$

where $c(t_i)$ is the 5-minute average concentration for period i (i=1,12), and f is a function of the average concentration and which period is being modeled. One would apply mass balance constraints on the function and that it provides the observed distribution of concentrations. f(…) would likely give something approaching a log-normal distribution, where the geometric standard deviation is a function of the observed concentration (or maybe not, for simplicity: I worry that we might be trying to make this more complex than need be). I would have little problem (others might) that if the limited data you have can be reasonably fit with a log-normal distribution and that the gsd does not change that much between locations, use the resulting analytical form for a national assessment. This might provide a more efficient approach, and given that it would provide a more statistically founded concentration structure, could arguably be better.

Page 11, last line: I think you mean reporting, not measuring.

In replying to the given questions:

1. Is a proportional approach to just meeting more stringent standards reasonable?
    a. **Answer:** Yes, with a condition. The PRB is quite low, but it is likely that the application of controls to meet a more stringent standard will concentrate on large point sources, and very possibly ones near to the urban area. Thus, PRB may not be the appropriate point to which to extrapolate. Instead, a low value resulting from a distributed set of small ubiquitous sources may be better. Probably makes little difference, so I am not sure it is worth the effort to figure out what alternative vaule to use, but you might just mention this issue. (p.s., sorry about dickering about the difference between linear and proportional.)
2. Should one extrapolate up to just meeting the current standard or use historic air quality?
    a. **Answer:** I would choose using historical data, though testing the difference at one or two locations. For one, this will help lend confidence to how you extrapolate to meeting more stringent standards using the linear rollback to the ubiquitous background concentration discussed above.

<u>Exposure Analysis</u>:

1. In considering the exposure analysis broadly:
    a. Do Panel members have any comments on the general structure and overall three-tier approach that staff plans to use for the exposure analysis? Are the criteria that staff plans to use for deciding whether to conduct a Tier II or Tier III analysis clear and appropriate?

Answer: Yes. (The approach is fine.) EPA should compare and contrast their approach to that used for other pollutants, and document why different methods are used. Again, use each review to make the exposure and risk assessment a more systematic, documented and turn-key. One could see that in about three years (a couple more pollutants) that a system much like that used for air quality modeling is used such that with relatively little effort exposures, risks, variabilities, sensitivities and uncertainties can be calculated, and the system as a whole has been intensely reviewed such that staff need not spend such effort, and the community is more comfortable with the results.

b. Have the most important factors influencing exposure to SO_2 been clearly accounted for and described?

Answer: I would tend to say yes, they have been discussed, but not accounted for in the analyses. As noted above, the potential for mischaracterizing SO2 levels, and area-wide PMRs as they relate to exposures, is much higher for SO2 than other pollutants.

c. The draft plan describes the basis for and selection of population groups of interest (i.e., children, asthmatics (children and adults), and the elderly) for which SO_2 exposure estimates are to be developed. Do Panel members generally agree with the groups of interest identified in the draft plan?

Answer: Yes.

2. In considering the Tier I exposure assessment:

a. Do Panel members agree a statistical model using available ambient 5-minute data is appropriate…?

Answer: It is a fine start. One could conceive of a more comprehensive approach, but I am not sure it is worth the effort (see above).

b. Do Panel members agree with applying a statistical model to estimate 5-minute exceedences exist where only 1-hr monitoring is performed?

Answer: Mostly yes, but the uncertainty and sensitivity analyses should consider this source of uncertainty and potential bias. Further, the model to be used needs to be better explained and tested.

3. In considering a potential Tier II exposure assessment:

a. Do Panel members agree with the combined emissions/dispersion modeling approach to estimate short-term (hourly) SO_2 concentrations?

Answer: Yes, as long as the model is evaluated and performance documented.

b.

Answer:.

c. Do you agree with using PMRs to develop short term peak concentrations

Answer: See discussion above.

4. Is it reasonable to use APEX?

Answer: Yes, as long as the model is evaluated and performance documented.

5. Do Panel members have any comments or advice regarding the general approach to addressing uncertainty and variability in each Tier of the exposure assessment as described in the draft plan?

Answer: Provide, early on, results of some sensitivity analyses. Do not overestimate uncertainties going in.

Comments from Mr. Avol

SECTION 1

Pg 2, lines 5 and 6 – Won't the exposure and health risk assessments also consider longer-term exposure (such as one-hour or 24hr), and not just five minute exposures?

(Typo) P3, para2, line 4 – should be "…within a given day…"

(Typo) P3, para4, line 7 – should be "down", not "sown"

SECTION 2

(no comments)

SECTION 3

P8, last para – Given the importance of port-related SOx emissions (primarily from bunker fuel or distillates in ship main and auxiliary engines), should identify ships and ports as area worthy of closer scrutiny.

P15, bullet 2, first sentence – assertion that current ambient monitoring siting captures anticipated occurrences of five-minute SO2 peaks assumes emissions are coming from electrical utilities and conventional historic sources…In terms of 5-minute averages, drifting plume touchdown points can become important; what evidence is there that current monitors are properly sited for this? What evidence is there that current monitors are properly sited to capture 5-minutes peaks downwind of ports and shipping lanes in coastal areas (both urban and rural)?

P24, set of bullets – will these population sub-groups of concern be directly drawn from the revised ISA, or will these four groups be the only ones chosen (in other words, what about other susceptible subgroups)?

SECTION 4

P28, para1, (4) – How does having time and resources to complete under the schedule enter into this – isn't there a court-agreed-upon schedule to do it?

P29, para3, line4 and line9 – The assertion that the ISA found SO2 to be robust, generally robust, after adjustment for PM and other co-pollutants should be re-visited and clarified, based on the recent ISA discussions.

(Typo) p29, para3, line 11 – "…in both the and human…" has a word missing or an unnecessary "and"

P29, para3, last sentence – Based on John Balmes' comments regarding the ISA, this conclusion about SO2-induced altered lung host defense is incorrect.

P29, para 3 and para 4 – Given the discussion regarding clearer objective and consistent definitions of the causal chain criteria, these comments about "suggestive", "plausible", "inconclusive", etc should be re-visited following revision of the ISA.

P30, para3, last sentence – is this comment about not focusing on long-term exposure-related health effects consistent with the Agency's historical perspective on SO2? Didn't the previous review decline to establish a standard for short-term exposures, in lieu of long-term exposure? Does this create a logical problem?

P31, para1, last sentence – doesn't the validity of this statement depend on the correlation between 5 minute readings and 1hr or 24hr values…which we apparently do not have thus far?

P36, last bullet – How can scheduling and resources be a viable rationale for not going forward, if the framework for the decision to go to Tier III is in place and the eligibility criteria are met? Isn't this being unresponsive?

P37, last sentence – This presumes that living close to fossil-fueled emission sources are the greatest concern, but these sources have been generally well-controlled. How about living close to a port or next to a rail line (burning fuel with several thousand ppm S, if it is outside of California), or living along the coast by a ship transit corridor?

Comments from Dr. Samet

General Comments:

In general, this document satisfactorily describes the approach that will be taken to carry out the exposure assessment and risk assessment for SO_2. Initial links have been made to the draft ISA. The document states explicitly that it will focus on the health effects of SO_2 and not consider secondarily formed particulate species. This makes the clinical studies directly relevant to the risk assessment while implying that the risk estimates from the epidemiological studies should be adjusted for PM; consideration needs to be given to potential "double counting" of outcome events across PM and SO_2 risk assessments.

Exposure Assessment:

The exposure assessment is quite detailed in describing the planned approach to addressing the characterization of the frequency of short-term peaks and of human exposures, including relevant subpopulations to these peaks. This aspect of the analytic plan is well developed. I have the following specific comments:

- Given the possibility that the risk assessment may consider effects identified in epidemiological studies, should the plan be expanded to include longer-time frame concentrations and exposures?

- The uncertainty classification (given on page 7) needs development. What are the criteria for "minimal, moderate, and major" levels of uncertainty?

- Only a limited set of monitors offer information on 5-minute concentrations. Analyses directed at representativeness are needed.

- Can the exposure model incorporate the possibility that susceptible persons take behavioral measures to reduce exposures?

Risk Assessment:

The document describes a three-tiered approach to the risk assessment. General criteria are offered for determining whether to undertake a Tier III assessment; I recommend bringing greater specifity to the description of these criteria on page 28. Specific comments are given below:

- The third paragraph on page 29 is confusing.

- Page 31 describes the approach to be followed for the acute responses, based on the clinical studies, and incorporating the "benchmark concentration". The text should be made more explicit. Is the scaling of risk linear in concentration?

- Page 35 states that "…staff judges that a C-R function estimated in the assessment location is preferable to a function estimated in some other location,…" This judgment warrants reassessment, particularly given the statistical instability of single city estimates. For those health outcomes with risk estimates available

from multiple locations, the aggregate estimate seems preferable, unless there are particular characteristics of the location that would indicate that a locally derived estimate is preferable.

Comments from Dr. Ultmann

Air Quality Considerations

1. Proportional approach seems reasonable to roll back current standard.
2. Proportional approach is more reasonable than historical approach, because of changes in conditions such as source distributions that could affect 1-hour mean and peak exposures in different ways.

Exposure Analysis

1. Considering the exposure analysis broadly, the two-tier approach is prudent and adequately explained in the document.
2. To the extent that quantity (and quantity) of the data is sufficient to provide reasonable power to the analysis, the application of a statistical model is a logical plan.
3. Considering a potential tier II exposure assessment, I am in general agreement with the approach, but because of the numerous uncertainties in dispersion and APEX models, I encourage that an attempt be made to validate the model simulations. The use of the PMR CDFS is reasonable given the low ambient SO_2 levels and negligible background levels that would tend to linearize the values of the peak and the mean value.
4. I have no comments regarding the general approach to uncertainty.

Comments from Dr. Pinkerton

The introduction, overview and background on SO_2 NAAQS are extremely well written and excellent in providing a framework for this document. These materials provide a clear and necessary foundation to understand the basis for the scope and methods for implementation of exposure and risk assessment for sulfur dioxide.

The organization and flow of the chapters and content for the Scope and Methods for Exposure Assessment and Risk Assessment Plan are extremely lucid and well written with excellent figures and tables. The methods used for estimating 5-minute peak concentrations are clearly explained. The APEX approach to simulate exposures that occur in indoor, outdoor and in-vehicle microenvironments seems highly reasonable. The selection of populations modeled to include 1) children (birth to 18 years), 2) asthmatic children (birth to 18 years), 3) asthmatic adults and 4) elderly (greater than 65 years) is also highly appropriate.

Air quality considerations for this document are particularly challenging in view of the need to consider spatial and temporal levels of SO_2. A two-tiered approach as outlined should be rigorously followed and fully implemented, with an adequate degree of liberal application. Such an approach should provide the necessary tools to provide for the proper interpretation and analysis of exposure scenarios that could be associated with possible adverse health effects to better protect the public in the eventual decision to be made for the next SO_2 NAAQS.

The susceptible/sensitive populations have been adequately identified. However, due to the type of exposure to SO_2 involves in large measure in proximity to power plants, consideration should also be applied to exposure by occupation associated with this gaseous pollutant.

Annual and 24-averages from 1990 to 2006 clearly show a downward trend for SO_2. Those exposed to the highest concentrations of SO_2 are limited select areas such as the Ohio River Valley, the US Northwest and Hawaii. Therefore, these areas require the greatest consideration for health protection, rather than an average over multiple other areas experiencing relatively lower levels of SO_2.

Air quality considerations: In view of the anticipated low exposure levels, the use of a proportional approach to adjusting air quality to simulate just meeting more stringent alternative air quality seems to be a highly logical approach.

The discussion on variability and uncertainty is helpful to place into perspective variables across time and space for individual exposure as well as assessment to estimate the number of exceedances of alternative health effect benchmarks.

Minor comment:

Page 12: Figure 4 is referred to in the text as Figure 3.

Comments from Dr. Sheppard

Here is my summary understanding of chapters 3 and 4:

- Tier 1 exposure assessment is an exposure characterization aimed at predicting ambient concentration at the 5-minute level for use in a risk assessment based on evidence from controlled clinical studies.
- Tier II exposure assessment is an exposure characterization aimed at predicting personal exposure at the 5-minute level for use in a Tier III risk assessment. It includes a dispersion model of concentration, but this is aimed at the 5-minute predictions needed for clinical studies rather than usual population exposure needed for epidemiological studies.
- Tier I risk assessment is a qualitative evaluation of health risks that provides preliminary input into the Tiers II and III risk assessments.
- Tier II risk assessment for controlled clinical study outcomes is a quantification of exposure aimed at determining the number of times the population would be exposed to benchmark doses.
- Tier II risk assessment for epidemiological study outcomes is an evaluation of exposure patterns to help with determination of population-level health effects of SO2.
- Tier III risk assessment for controlled clinical study outcomes were not discussed in chapter 4, but if done, based on information in chapter 3 they will use APEX predictions of 5-minute personal exposures.
- If done, Tier III risk assessment for epidemiological study outcomes will use concentration data from specific locations, incidence data estimated from the same locations, and concentration-response functions from single- and multi-pollutant models.

Discussion:

- The goals of the exposure and risk assessment aren't well aligned. The exposure assessment focuses on exposures needed for controlled clinical studies while the bulk of the health assessment focuses on the epidemiological studies.
- Tier I exposure:
    - Clarify the definition and discuss the prior analyses of PMR to clarify how these were done. Define the term "peak" – I assume this is the maximum in the hour. Define whether a single random 5-minute measurement, the maximum 5-minute measurement, or all 12 5-minute measurements in each hour are used in the analyses presented and will be used for modeling planned. The data in Figure 3 shows many PMR values below 1 suggesting the 5-minute data are not actually peak measurements. Figure 4 suggests a different set of data were used.
    - Using different PMR distributions for different influential features is valuable. Both mean 1-hour concentration and proximity to source may contribute to the distribution of the maximum 5-minute concentration in an hour.
    - Clarify the notation and wording in 3.2.1.4 as it is confusing. (Is the model for an expected number (wording) or a probability(notation)?

Incorporate the conditioning on influential features in the notation. Is r continuous? Why is P=1-p a cumulative probability and conditional on m but p isn't?

- I find some of the statements in 3.2.3 hard to believe. (Perhaps it is just semantics?) Quality of the monitoring data doesn't only depend upon quality of data collection, but also on the alignment of the data with the intended use. In this case monitor siting features will affect the usefulness of the data. Temporal variability will also depend importantly on monitor siting.
- Looking ahead to the need for predictions of multiple measurements in an hour for use in Tier II, I question whether the PMR approach is appropriate. Wouldn't it be more appropriate to develop the distribution of 5-minute data conditional on the hourly mean? This would then allow all the values of interest to be simulated directly from the distribution. (refer to section 3.3.2)
- P 14: Monitors are so sparsely sited that results of the analysis will need to be applied to unmonitored areas. This is in contrast to the described approach which only addresses populations that live in the vicinity of existing monitors.
- p 15 monitor siting: Analyses recommended for the ISA and needed for modeling 5-minute data should inform this question and bound the uncertainty. The key issue is how population exposure is represented by reliance on existing monitors. It may be possible to make different modeling assumptions that directly take into account monitor siting.
- p 15 temporal representativeness: Clarify what the temporal profiles will be assumed to be representative of. Must this be assumed, or can an analysis be done to verify this assumption?
- p 15-16 spatial representativeness: Once again this is a feature that should be able to be characterized rather than assumed.
- p 16 monitor to exposure representativeness: This bullet is confusing as written. This topic is not very clear in the ISA; improved ISA analysis and reporting may help with this HAP topic as well.

- Tier II exposure:
  - Section 3.3.2 would not be so convoluted if the distribution of the 5-minute data were modeled given the 1-hour mean instead of modeling the PMR. Given the purposes of the modeling, it may not be important to take into account correlation of 5-minute measurements within the hour. This will certainly simplify the modeling.
  - The description of the APEX model is good and EPA has had good success using this model for ozone health assessment. Refinements needed for SO2 modeling are new challenges.
  - The combined uncertainty and variability analysis for APEX is commendable.
  - In assessing APEX, direct comparison of predictions to measurements will be difficult because measurements are only for ambient concentrations while predictions will be for personal exposures.

- Tier I risk:
    - I appreciate the explicit inclusion of qualitative risk assessment as its own tier. This makes the entire process more open and transparent.
- Tier II risk:
    - Separate 4.3.1 into separate subsections for controlled and epi studies.
    - I don't understand the distinction being made between the national-scale analyses and urban-oriented analyses. (para 2 4.3.1)
    - The evaluation of patterns of air quality data approach is ambiguous. This does not appear to be a risk assessment.
    - Section 4.3.3 introduces a second level of tier II risk assessment in a section discussing variability and uncertainty.
- Tier III risk:
    - Based on the Tier II exposure discussion, the conclusion that there is insufficient information to develop the Tier III risk assessment for human clinical studies does not appear appropriate.
    - Use the section heading to clarify the description is for the epidemiological study results.
    - I don't find differences in location of health studies versus risk assessment a compelling enough reason to not conduct the risk assessment.
    - Section 4.4.3: I suggest moving towards the unified uncertainty analysis approach described for exposure modeling in section 3.3.6.
- Determination of approach: If at all possible, decision-making should be driven by scientific considerations as opposed to time and resources.

Other comments:

- p. 7: Regarding the choice of how many tiers will be done in the exposure and health assessments, decisions need to be open and transparent. I suggest all tiers be addressed in the final document even if no analysis is done.
- P. 14: This approach will be used to simulate the probability of a short-term peak, not a realized value of a peak. Correct?
- P 19 before 3.3.2: Make sure to clarify that output are predictions. Insert the word "predicted" before "SO2 concentrations".

Response to charge questions:

Air quality considerations:

2. Recognizing that current ambient air quality concentrations are lower than the current standards, the draft Health Assessment Plan discusses two alternative approaches to simulating ambient SO_2 levels associated with just meeting the current SO_2 standards: use of historical air quality data (e.g., possibly pre-2000) when ambient levels were at or above the current standards, or use of a proportional (i.e., linear) approach to adjust SO_2 levels upward. Do the Panel members have advice or comments on these two alternative approaches to simulating air quality just meeting the current SO_2 standards?
I would use historical data. I suggest limiting how far back to go in time, even if this means using data that are below the current standard.

Exposure Analysis:

1. In considering the exposure analysis broadly:

a. Do Panel members have any comments on the general structure and overall two-tier approach that staff plans to use for the exposure analysis? Are the criteria that staff plans to use for deciding whether to conduct a Tier II analysis clear and appropriate?

Is the exposure assessment an analysis that should be done in isolation of the risk assessment? The planned uses of the analysis should be added to the criteria to conduct the analysis. Depending upon intended use, this could suggest a different conclusion than the other criteria, either less exposure analysis (e.g. if no risk assessment will rely on the data) or more exposure analysis (e.g. if a population summary of the numbers of people expected to experience certain exposures under certain activity conditions is needed).

b. Have the most important factors influencing exposure to SO_2 been clearly accounted for and described?

Should wind direction be incorporated?

c. The draft plan describes the basis for and selection of population groups of interest (i.e., children, asthmatics (children and adults), and the elderly) for which SO_2 exposure estimates are to be developed. Do Panel members generally agree with the groups of interest identified in the draft plan?

Yes. If possible the subgroup of responders should be represented. However it is not currently known how to identify this subgroup.

2. In considering the Tier I exposure assessment:

a. Do Panel members agree that a statistical model using available ambient 5-minute monitoring data is appropriate for estimating expected exceedances of very short-term (5-minute) potential health effect benchmarks?

Analysis needs to be done to determine the comparability of the monitors with all 12 5-minute averages in an hour vs. monitors with only the 5-minute maximum within an hour vs. the monitors with no 5-minute data within the hour. If for the purposes of this specific analysis, the data are comparable, or comparable subsets can be selected then I agree it is appropriate to use the available 5-minute data for developing the model. As noted above, I have questions about the notation in section 3.2.1.4., the basis for Figure 3, and the plan to rely on the peak to mean ratio as the quantity of interest. I suggest instead modeling the distribution of 5-minute averages given the hourly mean as an alternative to modeling the PMR.

b. Do Panel members agree with the approach of applying a statistical model to estimate 5-minute concentration exceedances at monitoring locations where only 1-hour monitoring was performed for evaluating the extent of 5-minute peaks associated with meeting alternative standards with longer averaging times?

No. See my comments above questioning the need to model the PMR.

3. In considering a potential Tier II exposure assessment:

a. Do Panel members agree with the combined emissions/dispersion modeling approach to estimate short-term (hourly) SO_2 concentrations in close proximity to SO_2 emission sources?

Seems reasonable

b. Do Panel members have comments or advice regarding the described binning of sources and development of prototype stacks/facilities?

Seems reasonable

c. Do Panel members agree with the approach using peak-to-mean ratio cumulative density functions (PMR CDFs) to estimate very short-term peak concentrations from the 1-hour modeled concentrations?

I question the decision to model the PMR, particularly when there is interest in multiple "peaks" within an hour. If peak means maximum, this concept doesn't even make sense.

d. Do Panel members generally agree that the approach described using APEX is reasonable and appropriate to estimate the occurrence of very short-term (5 minute) SO_2 peak exposures?

I think 5-minute concentrations given the hourly average should be estimated/predicted directly from the distribution of 5-minute data given the mean, rather than focusing on the peaks. Is it reasonable to assume target ventilation/activity levels would not vary within a person over time?

4. Do Panel members have any comments or advice regarding the general approach to addressing uncertainty and variability in each Tier of the exposure assessment as described in the draft plan?

The unified uncertainty analysis for Tier II is good and I agree with the plan to incorporate it into Tier II.

Health Risk Assessment:

1. Do Panel members have any comments on the general structure and overall three-tier approach that staff plans to use for the risk assessment? Are the criteria that staff plans to use for deciding whether to conduct a Tier III risk assessment clear and appropriate?

I like the explicit addition of the Tier I qualitative risk assessment. Criteria for judging information suitable for a Tier III risk assessment shouldn't be so strict that it limits the possibility of conducting a quantitative risk assessment. For instance, I think a criterion that the time series study estimates come from the same cities where they will be applied is too strict.

2. In considering the Tier I risk assessment:

a. Do Panel members agree with the approach of having a qualitative assessment of health endpoints to identify which are likely candidates for a more sophisticated and quantitative tier of assessment?

Yes. This section should also discuss endpoints that would inform a qualitative assessment but which would not progress to a quantitative analysis.

b. Do Panel members agree with our initial observation that controlled human exposure studies demonstrate strong evidence for bronchoconstriction in exercising asthmatics following 5-10 minutes SO_2 exposure?

Yes

c. Do Panel members agree with staff's initial observation that the strongest epidemiologic evidence is for respiratory symptoms in asthmatic children and respiratory-related hospital admissions and respiratory-related emergency department visits in asthmatics and others with respiratory conditions?

Yes

3. In considering the Tier II risk assessment:

a. In general, are staff plans to use potential health effect benchmarks to address respiratory effects demonstrated in exercising asthmatics in controlled human exposure studies clear and appropriate?
I don't understand why short-term peaks in the vicinity of SO2 sources won't be considered since this is expected to be the major source of high 5-minute exposures. (sec 4.3.1 end of paragraph 2)
b. Do Panel members generally agree with the tentatively identified potential health effect benchmark of 0.5 to 0.6 ppm for exercising asthmatics following 5-10 minutes SO_2 exposure?
The ISA mentioned there were responders down to 0.25 ppm.
c. Do Panel members generally agree with the staff's approach of focusing on areas around major sources of SO_2 with respect to concerns about 5-10 minute peak exposures related to the respiratory effects observed in controlled human exposure studies?
The wording in section 4.3.1 of the document suggests the opposite.
d. Do Panel members generally agree with staff's approach of focusing on urban areas with respect to concerns about 1- and 24-hr and annual SO_2 concentrations related to respiratory effects observed in epidemiologic studies?
Focus on urban areas for time series studies is appropriate, but I don't understand what analysis will be done or reported.
e. Do Panel members have any comments or advice with respect to staff's approach of gathering additional information to characterize the SO_2 ambient air quality that existed at the time various key U.S. and Canadian studies addressing respiratory effects were conducted to see if the concentration-response relationships observed in these epidemiologic studies are related to particular SO_2 levels and associated averaging times, geographic location and/or season, and the inclusion of various co-pollutants?
This section wasn't clear.
4. In considering a potential Tier III risk assessment:
a. Do Panel members generally agree that there is insufficient information to develop credible exposure-response relationships for use in a quantitative risk assessment based on the controlled human exposure evidence?
No
b. Do Panel members have any comments or advice with respect to the general approach or specific factors to be considered in deciding whether or not to proceed to a Tier III quantitative risk assessment for the respiratory-related health endpoints based on epidemiologic evidence discussed in the draft plan?
Decision to proceed with a Tier III risk assessment for epidemiological studies of SO2 hinges upon making the assumption that the concentration-response estimates from these studies reflect a direct effect of SO2 on respiratory outcomes. This assumption can be questioned based on one or more of: the ability to attribute the effect to SO2 given the complex nature of the pollution mixture, the often secondary role of SO2 in the study analyses, and the ever-present challenge of making causal inference from observational studies. Taking a precautionary approach, I suggest proceeding with the quantitative estimates using both multi- and single-pollutant SO2 models as is planned, and with explicit recognition of the assumptions needed to even conduct the assessment.

5. Do Panel members have any comments or advice with respect to the general approach to addressing uncertainty and variability in each Tier of the risk assessment as described in the draft plan?

To the degree possible, move towards conducting a unified uncertainty analysis similar to the one described in the exposure assessment section.

Comments from Dr. Balmes

In general, I think the approaches described in the document are appropriate to provide important information that will be useful in the process of review of the NAAQS for SOx. I applaud the agency for the estimation of 5-min peak exposure data. This is a relevant exposure metric given the consistent results from controlled human exposure studies showing bronchoconstriction with such short duration exposures. I also am pleased that the agency is considering the use of emergency department (ED) asthma visit data in the risk assessment. This is especially important given the relative weight of the epidemiologic evidence of an association between SO2 exposure and asthma outcomes, especially in children. Asthma hospitalization data only capture the tip of the iceberg of asthma morbidity and the addition of ED data allow a greater proportion of the burden of asthma to be assessed. I made this same point during the ozone NAAQS review and was told that for the cities in which the exposure assessment was done, there were inadequate asthma ED data to use in the risk assessment. For this review, more ED data should be available – for example, California now has ED as well as hospital discharge data available state-wide.