Preliminary Comments from Members of the CASAC Sulfur Oxides Panel on EPA's Integrated Science Assessment for Sulfur Oxides – Health Criteria (Second External Review Draft – December 2016)

Received as of 03-16-17

Dr. Judith Chow 2
Dr. Alison Cullen 5
Dr. Delbert Eatough 9
Dr. H. Christopher Frey 18
Dr. Frank Speizer 29
Dr. James Ultman 38

Dr. Judith Chow

Comments on Chapter 2

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

Section 2.2 Anthropogenic and Natural Sources of Sulfur Dioxide U.S. Anthropogenic Sources: SO_2 emissions are a rapidly moving target, especially for electric generating units (EGUs), as illustrated in the attached Figure 1 (a version of which should replace Figure 2-5 [Page2-10] as more relevant to the current situation). The attached Table 1 shows substantial differences relative to the ISA's Table 2-1 (Page 2-9). A modified Figure 2-5 might also include the 2006 on-road and 2007 off-road diesel sulfur fuel standards.

Section 2.5 Environmental Concentrations
Median (50th percentile) concentrations in Figure 2-23 for Pittsburgh (Page 2-58) is a straight line and nearly straight line (except for hour 3:00) for NY in Figure 2-24 (Page 2-59). Section 2.5.1 noted that negative values in the AQS database were excluded in the analysis but concentrations below the lower detection limits (LDLs) were included (Page 2-32, Lines 6-8). As SO_2 concentrations varied diurnally, why are these straight lines at the concentration level around 1 ppb, about 50% of the LDLs of 2 ppb? For Figures 2-23 to 2-25 (Pages 2-58 to 2-61), a footnote should be included that specify how the mean (or arithmetic average) was calculated when many measurements were below the LDLs.

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

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

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

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

References:

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

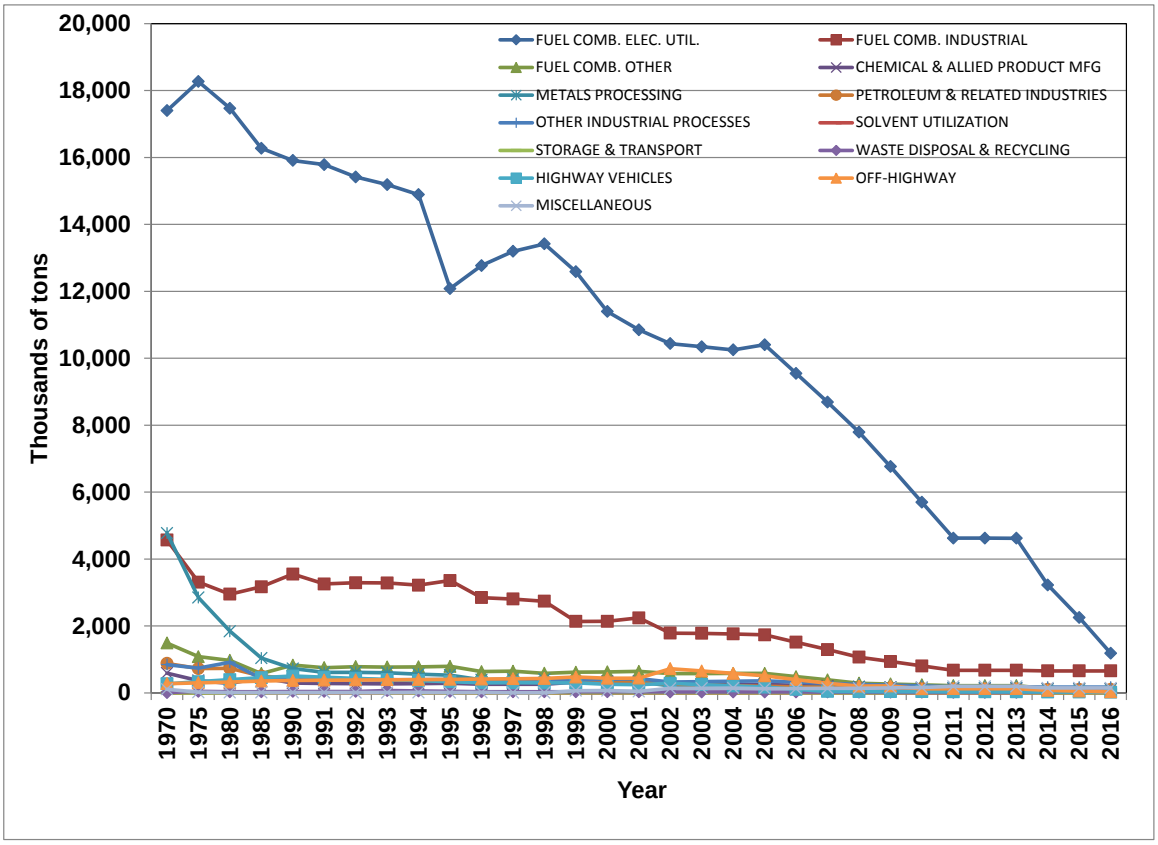

Figure 1. SO2 emission trends through 2016 from the spreadsheet available at https://www.epa.gov/air-emissions-inventories/air-pollutant-emissions-trends-data. The difference between ISA Figure 2-5 that ends in 2010 and the 2016 emission estimates for EGUs is substantial.

Table 1. Changes in U.S. SO_2 emissions since 2001.

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

Dr. Alison Cullen

Comments on Chapter 3 - Exposure to Ambient Sulfur Dioxide

EPA has strengthened this chapter since the last draft.

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

Specific points and comments are presented below.

Section 3.2.1 Exposure Metrics

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

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

Section 3.2.2 Conceptual Model of Personal Exposure

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

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

Section 3.3 Methodological Considerations for Use of Exposure Data

In the opening paragraph, the measurements of microenvironmental concentrations are clearly designated as surrogates for personal SO2 exposure. And in Section 3.3.2 the modeled concentrations are designated as contributing to surrogate estimates. The sentence that follows in line 33 could use clarification however as it mentions factors such as time-activity patterns

and indoor concentrations of SO2 in microenvironments that are incorporated in a portion of the modeling.

Section 3.4 Exposure Assessment, Error and Epidemiological Inference

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

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

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

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

Section 3.4.2.1 Activity Patterns

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

Section 3.4.2.2 Spatial Variability

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

Section 3.4.2.3 Temporal Variability

In the Guay et al (2011) study the population did not move often, an important factor in a long term analysis. It would be helpful to comment in the text on the applicability of this

characteristic of the Canadian population, as reflecting the distribution of the population between urban and rural settings, the difference between the US and Canada populations, etc.

Section 3.4.3 Copollutant Relationships

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Pg 3-52 line 12-13 rewrite for clarity

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

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

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

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

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

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

Dr. Delbert Eatough

Comments on Chapter 2

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

I focus on the following:

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

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

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

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

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

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

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

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

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

If the answer to 2. is yes, aerosol inorganic S(IV) does appear to be a significant confounder of SO_2 exposure, then the observations seen in 1. become important with respect to health criteria evaluations. I believe the question is significant enough that EPA should not just avoid seeking these answers as they have done in this Second Draft.
For the above outlined reasons, I first address the lack of response to the above outlined requests of CASAC by EPA in Chapter 2.

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

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

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

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

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

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

Figure 2-5. If iron and steel and smelters were included as specific sectors in this plot as requested, the reasonableness of looking at older smelter related epidemiological studies to determine if aerosol S(IV) species could be a confounder would be apparent. My suggestion in my final comments given above has been ignored. Why?

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

This source information could be included in a Table for the two highest concentration ranges.

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

the location of these highest sites and major nearby sources would also be valuable here and would respond to the request of CASAC.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

This question has been completely ignored in the ISA. In my previous final comments on the first draft ISA I included a detailed section, requested by other members of the committee, on our research in this area. While other aspects of both the formation of SO_2 in the atmosphere and the

conversion of SO_2 other species is well covered, the question of identifying non-sulfate S(IV) and S(VI) gas and particulate species from SO_2 formation is not touched on at all. The HERO summary of references both considered and used and considered and not used contains 16 of the references in my final comments on the last draft ISA. The only one used in the currently ISA was a paper on the rapid conversion of SO_2 in the plume from an oil-fired power plant in a fog bank, compared to when the plume was not in a fog bank. That is referenced in Chapter 2. That study also reported on the formation of gas phase dimethyl sulfate when the plume was not in the fog bank. That chemistry is not mentioned anywhere in Chapter 2.

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

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

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

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

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

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

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

I have outlined above the limited toxicological research (Amdur 1975, Colluci 1976) which indicates the combination of transition metal salts and SO_2 leads to a significant enhancement in the exacerbation of respiratory symptoms in experimental animals. I have also alluded to the research by our group on the development of a titration calorimetry oxidation methods (Hansen 1976) and the use of ESCA analysis (Eatough 1978) to study formation of transition metal S(IV)

species in emissions, particularly from smelters and integrated steel mills. Based on early results from the combination of these two observations it has been previously postulated (Colluci and Eatough, 1976)**:**

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

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

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

Other Comments.

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

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

How common really was this occurrence in the data set?

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

Comments on Chapter 1, Integrative Synthesis of the ISA.

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

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

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

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

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

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

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

References

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

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

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

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

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

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

Dr. H. Christopher Frey

Comments on Chapter 3

The main comments on this chapter are:

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

What follows are detailed point-by-point comments.

Page 3-2, line 36 and after, "Exposure Error" is defined as error with using "concentration metrics" to represent actual exposure. The definition here is unclear. What is an "exposure metric"? How is an exposure metric different from a "surrogate"? A conceptual diagram may

help to clarify these. Is "metric" meant to encompass either an unbiased estimate (e.g., personal exposure) or a biased estimate (e.g., surrogate such as ambient exposure at a central site monitor?). As a matter of terminology, not all monitors are centrally located. The term "fixed site monitor" may be more accurately and descriptive of a regulator ambient monitor at a fixed location.

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

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

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

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

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

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

Page 3-7, line 30. The header for this section is "Modeling," and later it is revealed that the scope of this section includes stochastic population-based exposure modeling (also referred to in

the ISA as microenvironmental modeling). However, the first sentence of this section is only applicable to models that predict outdoor concentration. As another example, the sentence starting on line 33 states that models do not estimate exposures directly, but in fact this is the purpose of stochastic population-based models. Such models account for time-location patterns and indoor concentrations in various microenvironments. Thus, there is a mismatch between this introduction and the content of the section. The introduction paragraph should be rewritten to more accurately introduce the scope of this section.

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

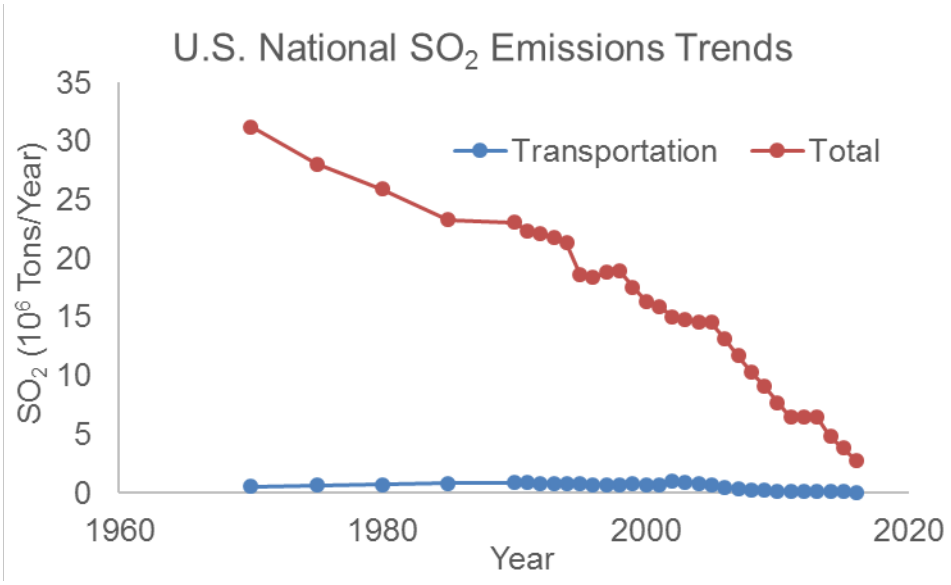

Figure 1. Annual National SO2 Emissions versus Year for Transportation and Total

Page 3-8, line 23. As a matter of context, it should be mentioned that motor vehicles contribute very little to the national SO_2 emission inventory. For example, as shown in Figure 1, emissions from onroad and nonroad transportation combined contribute an average of 3 percent to the total

U.S. national emissions since 1970, and in recent years (since 2010) have contributed only 2 percent. The highest annual contribution was 7 percent, in 2002. It can be the case that the relative contribution of transportation sources may be larger in a given region or urban area. The sentence on lines 22-25 is confusing. Why would a "decrease" in SO_2 be expected "near a highway" – did the authors mean to say that there was not a significant difference in SO_2 in proximity to a highway versus elsewhere in the study region?

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

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

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

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

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

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

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

The factors above are suggestions – this could be refined.

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

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

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

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

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

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

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

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

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

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

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

Table 3-1 is missing microenvironmental monitors.

Other comments on Table 3-1:

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

  - Frey, H.C., and Y. Zhao, "Quantification of Variability and Uncertainty for Air Toxic Emission Inventories With Censored Emission Factor Data," Environmental Science and Technology, 38(22):6094-6100 (2004).
  - Zhao, Y. and H.C. Frey, "Uncertainty for Data with Non-Detects: Air Toxic Emissions from Combustion," Human and Ecological Risk Assessment, 12(6):1171-1191 (Dec 2006)
- Should add a section on microenvironmental concentration monitoring, which can use instruments that are more precise and accurate than those used in personal monitoring.
- Source proximity model: a limitation to mention is that it does not account for emissions, stack parameters, plume dispersion, meteorology, or atmospheric chemistry. In stating limitations, it would help to identify limitations of one method that are addressed by another method in the same table. The discussion of error and uncertainty for SPM is hard to follow – there is too much detail here about "the circle formed….". Given that the SPM does not deal with structural aspects of the physics and chemistry of plume dispersion, the SPM is likely subject to bias moreso than imprecision. There may be imprecision associated with how the model is calibrated – e.g., in terms of a standard error of the estimate, but the bias is much more difficult to quantify.
- If ambient concentration predicted by the model is not the same as at a receptor, please distinguish between bias and imprecision. Bias would imply that the model systematically over- or under-estimates the true concentration. Imprecision would imply that there is random error in the estimate. Given that SPMs do not account for some of the key factors that affect ambient concentration, it seems more likely that bias is the more significant concern here. The model may also be imprecise. The text here is also unclear. Perhaps it is trying to compare the actual and predicted concentrations, but refers to ambient concentration at the receptor (is this the predicted or actual) and the average ambient concentration (where – at the receptor? And is this predicted or actual?).
- For emission weighted proximity models, mention that limitations include 'does not account for stack parameters, plume dispersion, meteorology, or atmospheric chemistry.' The text on errors and uncertainties should be revised – see comments on SPM. Similar to SPM, the text on errors and uncertainties goes into too much detail about the "circle" but not enough distinction of bias versus precision.
- Land use regression: in addition to "land use factors," LUR accounts for distance from a source or a surrogate of source strength. Limitations – to be consistent, the limitations should be stated as "does not account for emission rates, stack parameters, plume dispersion, or atmospheric chemistry, and may account for meteorology only in terms of wind speed or wind direction, depending on model formulation." There should be more discussion either in the text or here as to why spatial resolution affects bias – this is not

very clear. If a model is mis-specified, then it is biased. Why is imprecision of concern for a mis-specified model?

- Inverse distance weighting and kriging: see comments above with regard to how to organize some of this information. For example, for limitations, could state, in addition what is mentioned here, that it does not account for emissions, stack parameters, dispersion, and only accounts for meteorology and chemistry to the extent that it is calibrated to data that have similar meteorology and chemistry. With regard to bias and precision, if sources are not "captured" then it will be biased. Smoothing could also produce bias. While there is potential for negative bias, there is also potential for positive bias – e.g., if the concentration surface does not adequately account for SO_2 deposition or loss processes, or variation in atmospheric/meteorological conditions. Sources of imprecision should be mentioned that are distinct from sources of bias. What factors would lead to imprecision?
- For dispersion modeling, other strengths to mention are that they account for stack parameters, emission rates, mixing height, atmospheric stability, meteorology, terrain complexity.
- For Chemical Transport Models, strengths include accounting for emissions, mixing height, atmospheric stability, meteorology. The limitation is not quite right… local emission sources can be "accounted for" in a large grid cell in terms of being part of the inventory of emissions released into the grid cell, but the model is unable to estimate the ambient concentration associated with the plume from the local source, unless a plume-in-grid model is used.
- For the so-called "microenvironmental model" (which should be referred to really as a stochastic population-based exposure model), another application is to quantify inter-city variability in exposure estimates. The limitation is a bit confusing in that these models simulate synthetic individuals and can estimate inter-individual variability in exposures, but they do not represent actual, specific individuals in a population. The explanation given for errors and uncertainties is about bias, not "reduced precision".

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

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

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

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

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

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

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

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

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

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

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

Section 3.4.2.1: EPA did a nice job of updating and summarizing CHAD in the 2014 final Health Risk and Exposure Assessment for the Ozone review (EPA-452/R-14-004a). One issue that came up in that review was whether activity patterns might be changing over time,

especially for children. Since the SO2 review is likely to address similar subpopulations (e.g., school children with asthma), any updates to the 2014 HREA with respect to CHAD would be important to mention in this section. If there are not updates, then it would still be useful to reference the 2014 Ozone HREA and explain (briefly) the key findings that continue to be relevant here.

Are activity patterns different by location/region?

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Dr. Frank Speizer

Please note that comments that are in bold and italicized are for the most part direct quotes from the document. and are included as reminders to me to be used during the review process.

Comments on Executive Summary and Chapter 1

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

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

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

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

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

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

Comments on Chapter 3

Tables 3-3 and 3-4. These are important examples of the kinds of information that need to be combined with the data in Chapter 6 to make national estimates of populations at risk (see concerns on the way Chapter 6 is organized).

Figure 3.3 provides important evidence of high correlation between any of the standard measures up to 24 hours, with some outliers that will have to be considered. This will need to be considered in assigning a frequency of excessively or even modestly high levels.

Chapter 3 Summary: Maybe this needs to be in Chapter 4 but I am concerned that there are two issues falling between the cracks. 1). The level of correlation between SO2 and other NAAQS as demonstrated in Figure 3-3 on average is reassuring in dealing with the co-pollutant issues, however, much of the data making up these figures are derived for relatively low levels of exposure (Table 2.6, except for 1 hour max 90% of all values below 4 ppb) and as soon as even modestly higher levels occur the correlations jump above 0.6. 2) Although this is a chapter on exposure, there is an issue with regard to how this exposure gets to the airways. This is considered in Chapter 4, but the figures used in Chapter 4 assumes a "magical" transfer of SO2 exposure to some chemistry in the airway (see Figures 4.2, 4.3) --Perhaps this is discussed better in Chapter 2

Comments on Chapter 4

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

Page 4-13. In summary, inhaled SO2 is readily dissolved in the ELF where it exists as a mixture of bisulfite and sulfite with the latter predominating. Bisulfite reacts with disulfide groups forming S-sulfonates; sulfite can diffuse across cell membranes and reach the circulation. Following absorption in the respiratory tract, SO2-derived products (e.g., sulfite and/or S-sulfonates) are widely distributed throughout the body and have been observed in the blood and urine within 5 minutes of starting an SO2 exposure of surgically isolated nasal airways

4.14 In summary, the primary route of sulfite metabolism is by sulfite oxidase-catalyzed oxidation into sulfate. The sulfite oxidase levels vary widely among tissues with very low levels found in the lung and high levels found in the liver, which plays a major role in the 9 detoxification of circulating sulfite. Sulfite oxidase activity is also highly variable among species with liver sulfite oxidase activity in rats being 10−20 times greater than in 21 humans

***4-22 In summary, SO2 acts as both a sensory and a pulmonary irritant through activation of sensory nerves in the respiratory tract resulting in neural reflex responses*. Pulmonary irritant responses due to SO2 exposure result in reflex bronchoconstriction, especially in adults with asthma. Both cholinergic parasympathetic pathways involving the vagus nerve and inflammation contribute to reflex bronchoconstriction in asthmatic individuals.**

4-*28 In summary, a growing body of evidence supports a role for SO2 in exacerbating AHR and/or allergic inflammation in animal models of allergic airway disease, as well as in asthmatic individuals. In addition, repeated or prolonged exposure to SO2 promotes allergic sensitization in naive newborn animals*

4-29 . In summary, exposure to SO2 may result in effects outside the respiratory tract via activation of sensory nerves in the respiratory tract resulting in a neural reflex response or mediated by circulating sulfite.

Comments on Chapter 5

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Page 82, line 20-21: Again the word "imprecise" I understand that the author is looking for a word to indicate that the Confidence Interval crosses 1.0, but I am not sure this is the best way to convey it. In this case [−0.82% (95% CI: −1.9, 0.31) per 10-ppb 21 increase in 2-h avg SO2] the suggestion is that the effect is really null or at best a modest negative association that is not significant. I don't think the word "imprecise" conveys that. Note in other places (eg. page 87the words used for similar statistics are "limited evidence", page 89: "large 26 uncertainty estimates") ***This may be an issue that the reviewers should discuss.***

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

Page 91, line 9-1. I take some issue with the first half of this statement. The fact that all the studies do not point in the same direction might lead one to say "inconsistent" However, taken

together one could argue that there is reasonable biologic and epidemiologic plausibility of an increase risk. I would agree with second half of the effects of attenuation by PM.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Page 185 Given the limited epidemiologic evidence, the association between ambient 10 SO2 concentrations and venous thromboembolism is unclear

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Starting page 5-259, Table 5-39. There are 3 entries that do not give levels of exposure. At least for the Moogavkar one would have thought the range would be the same as for Dominicii as the

same data base is being used. I could not get to the Bellini or Atkinson papers, to details of the aerometrics. I find it hard to believe the MISA2 paper did not have them for SO2. In any case without some estimate of exposure hard to understand how the calculations for these papers is made in Figure 5-17 and 5-18

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

Page 278 In conclusion, the consistent positive associations observed across various multicity studies is limited by the uncertainty due to whether SO2 is independently associated with total mortality, the representativeness of monitors and the 24-h avg SO2 exposure metric in capturing the spatial and temporal variability in exposure to SO2 (Section 3.4.2.2 and Section 3.4.2.3), and the uncertainty in the biological mechanism that could lead to SO2-induced mortality (Section 4.3). Collectively, this body of evidence is suggestive, but not sufficient to conclude there is a causal relationship between short-term SO2 exposure and total mortality. Page 293 The overall evidence is inadequate to infer a causal relationship between long-term exposure to SO2 and total mortality among adults.

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

Comments on Chapter 6

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

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

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

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

Page 6-16. In conclusion, evidence is adequate to conclude that people with asthma are at increased 29 risk for SO2-related health effects. Asthma prevalence in the U.S. is approximately 8−11% across age groups (Blackwell et al., 2014; Bloom et al., 2013), and thus, represents a substantial fraction of the population that may be at risk for respiratory effects related to ambient SO2 concentrations.

Need to Discuss: My problem with the conclusion is that no real estimate of population at risk is made for any of the potential risk groups and thus much of the chapter is really a rehash of data in Chapter 5. In general I am not happy with this Chapter. It seems that much of what is present is simply an alternative form of the material in Chapter 5. To effectively really produce Populations at potential risk some estimates of actual number that might be affect are necessary. The problem here is that simply limited the discussion to the respiratory effect is simply summarizing the previous chapter. To make the effort worthwhile I would have thought that at risk estimate would have been made for either the whole population (which would be too large) of selectively for parts of the population for 10,20 50% of asthmatic kid with estimated excess ED visits. Frequency of hospitalizations for adults with COPD at risk of similar numbers as related to SO2 exposures. This may be what is and should be contained in the REA.

Dr. James Ultman

Comments on Chapter 4

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

It appears to me that the value of the Henry's law constant given in the ISA (pg 4-8, line 4) is actually the ratio of molar SO2 concentration in air to the total SO2 concentration in water, appearing as dissolved gas and as ions (Eq. 4-1). Strictly speaking, the Henry's law constant is the ratio of the SO2 partial pressure to the concentration (or mole fraction) of dissolved SO2 gas only. I suggest instead of Henry's law constant, the authors use another term, perhaps "solubility factor," for both SO2 with O3.

The SO2 mass transfer rates given on pages 4-9 (lines 3-4) and 4-10 (lines 6-12) seem to be based on computations made by the authors of the ISA. The computations rely on an equation for mass transfer rate (Asgharian, Eq. 3) that requires the concentration of SO2 to be known at the respired gas-ELF liquid interface. It appears that this concentration has been neglected so that local absorption is proportional to the gas-phase Sherwood number=(airway diameter)(gas phase mass transfer coefficient)/(gas phase diffusion coefficient). This is misleading since it is more likely that a transport resistance modulated by diffusion-reaction processes in the ELF produces a non-zero interfacial concentration which opposed absorption. The revised ISA should include a more complete description of the diffusion and reaction processes that contribute to absorption. A justification should be provided for ignoring the influence of ELF diffusion-reaction effects in the ELF. By the way, the assumption of a zero interfacial concentration is not consistent with the occurrence of SO_2 desorption during expiration, which is asserted in other places in the chapter.