Responses to CASAC Questions on the Draft PM PA from Consultant Mr. John J. Jansen
October 6, 2019

Given the broad nature of the questions posed by the CASAC members, I read a good deal of the draft Policy Assessment (not chapter 4 and only skimmed chapter 5 and appendix D for personal interest). I have not had time to read the draft ISA. My more detailed responses are to Dr. Boylan's questions on Chapter 2 (my area of expertise). I also have more general observations touching on questions from Drs. Cox and Frampton based on my observation of the NAAQS and CASAC processes over the years.

Questions from Dr. Boylan

Chapter 2

As to the discussion of emission sources, the discussion is very summary in nature but reasonably complete. I believe EPA should provide some perspective on the magnitude of biogenic VOCs. While the magnitude of natural sources of SO2, NOx, and NH3 are small relative to anthropogenic sources, that is not the case for biogenic VOCs. I believe it would also be helpful for more information about the relative contribution of various source categories at the urban/regional vs. national scale (e.g., see p 2-5). Having similar observations/discussions of SO2, NOx, and NH3 could provide insight to those who will be implementing the standard. For example, I find it interesting that in the modeling to just attain the current and alternative standards, many areas controlled by the annual standard needed to resort to proportional scaling to achieve the goal (see page C-32), substantial SO2 plus NOx emissions reductions were not sufficient. It is not clear why EPA did not add the primary scenarios first before resorting to proportional scaling as they did for the areas in the Primary PM scenario. Assuming the models are correct, this could have important implications to the states in their attainment demonstrations. I recognize the purpose of the PA is on setting the NAAQS, not implementation. Nevertheless, highlighting these findings now can help with a step (i.e., providing implementation guidance) EPA is notoriously late at achieving.

The discussion on monitoring is reasonably complete for the compliance networks but is a bit weak on the "Additional PM Measurements and Metrics" (page 2-18). There is a wealth of data on PM, PM composition, and pollutant gases from research networks such as the SEARCH network which, although now shut down, can supply important insights for model development and application and measurement errors. There is no mention in this section of continuous NO3 and NH3, both of which existed in SEARCH and other networks.

The discussion of measurement correlations and trends is reasonably complete. The section on composition trends, focusing only on sulfate, leaves the misimpression that sulfate alone trended down. Nitrate and carbon also did so but to a smaller degree. A more complete summary here would help. Further, the discussion of figure 2-11 on page 2-25 appears incorrect. The data for the Southeast is not an outlier. The Northwest and SoCal are. Clarification is needed.

Unless I missed it, I saw no discussion of the errors or biases in the measurements data generally nor the FRM/FEM specifically. Such errors and biases certainly exist, can be known, and would have some effect on the assessments. The magnitude of the effect would be study specific and, I would hope, be assessed by the authors of the health effects studies or risk assessors. I know that Dr. Paige Tolbert and the other members of the ARIES team took the issue seriously. As to the PM PA, I do not see that EPA has done a quantitative uncertainty analysis and, thus, has not included this issue. I am certainly a

proponent of a quantitative uncertainty analysis being performed. And while I cannot respond with specifics to **Dr. Lange's question number 4**, I do believe substantial data does exist to derive estimates for many of the items in her Table 1 and EPA should get on with performing that work before the next review. They have been advised to do so in the past.

The hybrid modeling is becoming a more common tool for estimating exposures in health studies. And there have been several studies looking at their performance. And while the discussion in the PA seems reasonable, it needs to carry the evaluation through the health assessments. On page C-4, EPA places a priority on selecting studies that indeed use multiple hybrid modeling approaches and compare the health results. However, Table C-1 lists only 1 such study, Jerrett, et. al., 2016. However, I could find no discussion in the PA nor the draft ISA of this feature of the study. What were the results of the comparison? It seems that at a minimum, EPA would discuss these results to support their claim on page 2-48 that hybrid modeling results "are reliable for estimating PM2.5 exposure in many applications" and "also suggests that the fields are reliable for use in health studies."

As to model performance, as with all regional model performance approaches, average operational performance is the norm. Gone are the days when extensive diagnostic performance was included. Recognizing that such performance evaluation is not going to happen, at least some look at the range of performance across locations would be helpful. What is the range of overestimation? Under estimation? Is the range better or worse in populated areas than in unpopulated areas? On page C-23, Table C-6 model performance (not hybrid) shows over prediction in the east and under prediction in the west, in some cases by as much as 2 ug/m3. Knowing such information would help with interpreting the adequacy of the hybrid models for health purposes.

The discussion of background needs clarification. For example, statements on pages 2-49 (lines 28 to 32) and 2-52 (lines 14 to 16) leave the role of non-US anthropogenic emissions unclear given the examples shown. The discussion of such emissions on page 2-53 suggests they are small due to short residence times of PM and yet natural sources appear more prominent (see next section). Shouldn't the residence time argument apply to all PM, regardless of source? On page 2-54 EPA claims that the remaining sulfate at the remote sites is from domestic SO2 emissions. Why?

Chapter 3

See below for a more general response to the adequacy of EPA's evidence and risk based approaches to reviewing the primary PM2.5 standard. The criteria for selecting the study areas seem reasonable but entire central portion of the country is missing, why? EPA provides a brief rationale why it decided to focus only on all-cause mortality on page C-2. The explanation implies morbidity outcomes would be less reliable? Was time, resources, and/or insufficient "data" also a factor? EPA's criteria for choosing the concentration-response functions is provided in section C.1.1. I have always had concerns that studies for cancer do not adequately consider the long latency and lack knowledge of historical exposures and population migration. I do not agree with EPA's exclusion of "groups not differentiated by socio-economic . . . attributes." Studies that fail to account for socioeconomic status are missing a key factor affecting health. Seeking studies that use multiple approaches to estimating exposures and calculating health risks from each is laudable. Unfortunately, only one such study was included (Jerrett, et. al., 2016) and EPA fails to describe these results nor use them in the uncertainty discussion. On pages C-22 and C-23 EPA is asserting that an epidemiological study on a cause specific mortality provides biological plausibility for all-cause mortality (i.e., ischemic events and heart failure for cardiovascular mortality and COPD and asthma for respiratory mortality). I do not believe this is correct. Biological

plausibility is provided through human exposure and animal toxicology studies (see page C-25 for a such a discussion). I recommend EPA refrain from blurring the definition of biological plausibility. As to additional research needs, the current debate over causality determinations is a clear area for EPA to explore. Also the approach reported in the studies by Dr. Anne Smith in Risk Analysis (see references provided by consultant Dr. Warner North) should have already been pursued by EPA since the last review.

Appendix C

See my comments on Chapter 2 above and the general observations below for some responses to the methods used in the risk assessment. I will say that the methods have become more complex over the years, supposedly to improve the credibility of the estimates. So many steps are added based on assertions of adding accuracy but little quantitative demonstration is made. Are we really getting more accurate estimates of exposure?

Questions from Dr. Packham

I have provided research recommendations elsewhere. As to the last question, effects from strategies to implement the NAAQS, EPA and CASAC have traditionally avoided this mandate. EPA by not providing the necessary information for CASAC to review and CASAC by not having the expertise, through its members or expert panels, on such issues to provide advice. Similar to conducting new causality determination approaches or new risk assessment approaches (e.g., suggested by Dr. Anne Smith), EPA needs to take the issues seriously and undertake the work. I recognize that too often time and resources are limited but it is time to give these, and other, issues priority.

General Observations

The issue of causality determinations is currently a particularly contentious issue. EPA has updated and refined its approach to causality determinations over the years but it has always been qualitative in nature (i.e., Causal relationship: the pollutant has been shown to result in health effects at relevant exposures based on studies encompassing multiple lines of evidence and chance, confounding, and other biases can be ruled out with reasonable confidence.). Despite suggestions to alter and/or add to its approach (many to add quantification), EPA has continued to justify the status quo as consistent with past practices endorsed by CASAC.

I view the suggestions of Dr. Cox to be an attempt to add quantification with which I concur, but I am not a statistician. And I do not know whether applying these techniques will confirm causality or refute it. I will confess to believing that associations are not and cannot be causal. Something more is needed (e.g., human exposure and/or animal toxicological studies, more on that below). A key problem is that the work to apply these techniques has largely not been done. Therefore, such results are not available for this review. There is also, of course, the problem associated with applying the techniques to existing study data bases due to confidentiality issues. While the work suggested by Dr. Cox may be problematic in this review, EPA has failed to follow-up on past, certainly doable recommendations such as applying the risk assessment techniques recommended by Dr. Anne Smith during the last review or completing a quantitative uncertainty analysis. That leaves us with the status quo approaches and asking whether the evidence has changed significantly.

EPA has summarized the new studies, most conducted with concentration levels lower than before (certainly a change from the last review) and with new approaches to establishing exposures, and new approaches to conducting the risk assessment. EPA concludes that the new information reaffirms and strengthens previous conclusions of causality, that key uncertainties have been reduced, and a lowering of the annual standard is supported. Since I am not sure what strengthening means (more of the same?), I do not view it as justification to change the NAAQS. Looking at the reduction in key uncertainties, EPA finds no threshold (no change), the shape of the concentration response (C-R) curve at lower concentrations is linear to zero (no change), human exposure and animal toxicology studies (necessary for providing biological plausibility) are still conducted at concentrations above ambient (no change), and the uncertainties of the C-R has decreased at lower concentrations.

With respect to decreasing uncertainties in the C-R functions at lower concentrations, I ask whether this isn't expected. Since there is no threshold and the shape of the C-R is linear to zero, the uncertainty at the lower concentration levels must decrease as there is now more data at those levels. Does this really add to our confidence in the risk results? It is not apparent given the confidence intervals in the risk results (see for example table C-5 on page C-85). EPA does not describe what the ranges are (e.g., uncertainty and/or variability). Nevertheless, they are quite large and call into question whether the reduced uncertainty has led to a different result than last time.

On the issue of biological plausibility, I reiterate that a) this is necessary to combine with associations to infer causality, and b) EPA should refrain from using an association to infer biological plausibility of another association. I also offer a different perspective on the results from human exposure and animal toxicological studies summarized in Table 3-2 on page 3-46. While these results still may be at levels higher than ambient (but closer than in the past), the trend in the results is intriguing. Of the studies with the lowest concentrations one shows no effect and one shows mixed results. Even the next few higher concentration studies show mixed results. And all of these studies are looking at biological markers that "demonstrate short-term exposures to PM2.5 *may* result in the types of cardiovascular endpoints that *could* lead to emergency department visits and hospital admissions in *some* people." (emphasis added, page 3-28). One might infer from this perspective that there are no adverse effects at ambient concentrations.

Finally, there are the risk estimates. As mentioned above, the confidence intervals (or whatever they are) in the tables such as 3-5 on page 3-85 are large. And the means of the alternative standards are within the range of the current standard. So are the changes really different than the current standard?