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

**OFFICE OF THE ADMINISTRATOR**
**SCIENCE ADVISORY BOARD**

October 24, 2006

EPA-CASAC-07-001

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

Subject: Clean Air Scientific Advisory Committee's (CASAC) Peer Review of the Agency's 2nd Draft Ozone Staff Paper

Dear Administrator Johnson:

EPA is in the process of reviewing the national ambient air quality standards (NAAQS) for ozone ($O_3$) and related photochemical oxidants, which the Agency most recently revised in July 1997. As part of its ongoing review of the ozone NAAQS, EPA's Office of Air Quality Planning and Standards (OAQPS) developed a 2nd Draft Ozone Staff Paper, entitled, *Review of the National Ambient Air Quality Standards for Ozone: Policy Assessment of Scientific and Technical Information* (July 2006). At the request of the Agency, EPA's Clean Air Scientific Advisory Committee (CASAC or Committee), supplemented by subject-matter-expert panelists — collectively referred to as the CASAC Ozone Review Panel (Ozone Panel) — met in a public meeting in Durham, NC, on August 24-25, 2006, to conduct a peer review of this draft Ozone Staff Paper and three related draft technical support documents.

In its summary of EPA staff conclusions on the primary (health-related) ozone NAAQS found in Chapter 6 of the 2nd Draft Ozone Staff Paper, OAQPS set-forth two options with regard to revising the level and the form of the standard: (1) retain the current primary eight-hour (8-hr) NAAQS of 0.08 parts per million (ppm); or (2) consider a reduction in the level of the primary $O_3$ NAAQS within the range of alternative 8-hr standards included in Staff's exposure and risk assessments (which included a range from 0.064 to 0.084 ppm) with primary focus on an $O_3$ level of 0.07 ppm with a range of forms from third- through fifth-highest daily maximum. The Ozone Panel unanimously concludes that:

1. There is no scientific justification for retaining the current primary 8-hr NAAQS of 0.08 parts per million (ppm), and

2. The primary 8-hr NAAQS needs to be substantially reduced to protect human health, particularly in sensitive subpopulations.

Therefore, *the CASAC unanimously recommends a range of 0.060 to 0.070 ppm for the primary ozone NAAQS*. With regard to the secondary (welfare-related) ozone NAAQS, t*he Ozone Panel is in strong agreement* with the scientific and technical evidence presented in the summary of EPA staff conclusions on the secondary ozone NAAQS found in Chapter 8 of the draft Staff Paper *in support of the alternative secondary standard of cumulative form that extends over an entire growing season.*

The Ozone Panel members agree that this letter adequately represents their views. The chartered Clean Air Scientific Advisory Committee fully endorses the Panel's letter and hereby forwards it to you as the Committee's consensus report on this subject. A discussion of each chapter in the 2nd Draft Ozone Staff Paper follows this letter, and the comments of individual Panel members on the 2nd Draft Ozone Staff Paper and three related draft technical support documents are attached as Appendix D.

## 1. Background

Section 109(d)(1) of the CAA requires that the Agency periodically review and revise, as appropriate, the air quality criteria and the NAAQS for the "criteria" air pollutants, including ambient ozone. Pursuant to sections 108 and 109 of the Act, EPA is in the process of reviewing the ozone NAAQS. OAQPS, within the Office of Air and Radiation (OAR), developed the 2nd Draft Ozone Staff Paper as part of this activity. In February 2006, the Agency's National Center for Environmental Assessment, Research Triangle Park, NC (NCEA-RTP), within the Agency's Office of Research and Development (ORD), released its final *Air Quality Criteria for Ozone and Related Photochemical Oxidants*, *Volumes I, II, and III*, (EPA/600/R-05/004aF-cF, Final Ozone Air Quality Criteria Document) for this current review cycle for the ozone NAAQS. The 2nd Draft Ozone Staff Paper evaluates the policy implications of the key scientific and technical information contained in the Final Ozone AQCD and identifies critical elements that the Agency believes should be considered in its review of the ozone NAAQS. The Ozone Staff Paper is intended to "bridge the gap" between the scientific review contained in the Ozone AQCD and the public health and welfare policy judgments required of the EPA Administrator in reviewing the ozone NAAQS.

The Ozone Panel met in a public meeting on December 8, 2005 to conduct a consultation on EPA's 1st Draft Ozone Staff Paper and two related technical support documents. However, given that the OAQPS' first draft Staff Paper did not contain Agency staff conclusions about whether to retain or revise the existing primary and secondary Ozone standards, the CASAC's activity only amounted to a technical assessment of that document. The Committee's letter to you from that meeting (EPA-CASAC-CON-06-003), dated February 16, 2006, is posted at URL: http://www.epa.gov/sab/pdf/casac_con_06_003.pdf.

## 2. CASAC Ozone Review Panel's Peer Review of the 2nd Draft Ozone Staff Paper and Related Technical Support Documents

The Ozone Panel reviewed the 2nd Draft Ozone Staff Paper and found it improved over the earlier version that had been reviewed as part of a consultation process. *However, the Panel did not agree with the EPA staff conclusions that it was appropriate to consider retaining the current NAAQS as an option that would be protective of public health and welfare.* The Ozone Panel's recommendations for reducing the level of the primary ozone standard, and its rationale for these recommendations, are provided immediately below. Following a detailed discussion on the primary and secondary NAAQS are the Panel's major, chapter-specific comments. Finally, the individual written comments of Ozone Panel members on the 2nd Draft Ozone Staff Paper and the three related draft technical support documents are attached in Appendix D. Panelists' responses to the Agency's charge questions are included in these individual review comments.

### Primary Ozone NAAQS

*New evidence supports and build-upon key, health-related conclusions drawn in the 1997 Ozone NAAQS review.* Indeed, in the 2nd Draft Ozone Staff Paper, EPA staff themselves arrived at this same conclusion:

> "Based on the above considerations and findings from the [Final Ozone AQCD], while being mindful of important remaining uncertainties, staff concludes that the newly available information generally reinforces our judgments about causal relationships between $O_3$ exposure and respiratory effects observed in the last review and broadens the evidence of $O_3$ -related associations to include additional respiratory-related endpoints, newly identified cardiovascular-related health endpoints, and mortality. Newly available evidence also has identified increased susceptibility in people with asthma. While recognizing that important uncertainties and research questions remain, we also conclude that progress has been made since the last review in advancing our understanding of potential mechanisms by which ambient $O_3$, alone and in combination with other pollutants, is causally linked to a range of respiratory- and cardiovascular-related health endpoints." (Pages 6-6 and 6-7)

Several new single-city studies and large multi-city studies designed specifically to examine the effects of ozone and other pollutants on both morbidity and mortality have provided more evidence for adverse health effects at concentrations lower than the current standard. (See the numerous ozone epidemiological single-city studies shown in Figure 3-4 on page 3-53 of the 2nd Draft Staff Paper and, in addition, Appendix 3B of the staff paper, which contains the summary of effect estimates and air quality data for these studies and multi-city epidemiological studies.) These studies are backed-up by evidence from controlled human exposure studies that also suggest that the current primary ozone NAAQS is not adequate to protect human health (Adams, 2002; McDonnell, 1996).

Furthermore, we have evidence from recently reported controlled clinical studies of healthy adult human volunteers exposed for 6.6 hours to 0.08, 0.06, or 0.04 ppm ozone, or to filtered air alone during moderate exercise (Adams, 2006). Statistically-significant decrements in lung function were observed at the 0.08 ppm exposure level. Importantly, adverse lung function effects were also observed in some individuals at 0.06 ppm (Adams, 2006). *These*

*results indicate that the current ozone standard of 0.08 ppm is not sufficiently health-protective with an adequate margin of safety.* It should be noted these findings were observed in healthy volunteers; similar studies in sensitive groups such as asthmatics have yet to be conducted. However, people with asthma, and particularly children, have been found to be more sensitive and to experience larger decrements in lung function in response to ozone exposures than would healthy volunteers (Mortimer *et al.*, 2002).

Going beyond spirometric decrements, adverse health effects due to low-concentration exposure to ambient ozone (that is, below the current primary 8-hour NAAQS) found in the broad range of epidemiologic and controlled exposure studies cited above include: an increase in school absenteeism; increases in respiratory hospital emergency department visits among asthmatics and patients with other respiratory diseases; an increase in hospitalizations for respiratory illnesses; an increase in symptoms associated with adverse health effects, including chest tightness and medication usage; and an increase in mortality (non-accidental, cardiorespiratory deaths) reported at exposure levels well below the current standard. *The CASAC considers each of these findings to be an important indicator of adverse health effects.* As demonstrated in Chapter 5 of the 2nd Draft Ozone Staff Paper (specifically, Figures 5.5, 5.7, 5.8, and 5.9), a significant decrease in adverse effects due to ozone exposures can be achieved by lowering the exposure concentrations below the current standard, which is effectively 0.084 ppm. Beneficial effects in terms of reduction of adverse health effects were calculated to occur at the lowest concentration considered (*i.e.*, 0.064 ppm). (See also Figure 3-4, "Effect estimates (with 95% confidence intervals) for associations between short-term ozone exposure and respiratory health outcomes," on page 3-53.)

The justification provided in the 2nd Draft Ozone Staff Paper for retaining the current level of the primary ozone standard as an option for the Administrator was based on results of controlled human exposure studies measuring modest declines in $FEV_1$ after exposures to 0.08 ppm ozone. However, as stated in the Staff Paper (page 3-6), while average decrements in the $FEV_1$ were relatively small, 26% of the subjects had greater than 10% decrements, which can be clinically significant. Also, while measures of $FEV_1$ are quantitative and readily obtainable in humans, they are not the only measures — and perhaps not the most sensitive measures — of the adverse health effects induced by ozone exposure. As stated on page 6-32 of the Final Ozone AQCD, "Spirometric responses to ozone are independent from inflammatory responses and markers of epithelial injury (Balmes *et al.*, 1996; Bloomberg *et al.*, 1999; Hazucha *et al.*, 1996; Torres *et al.*, 1997). Significant inflammatory responses to ozone exposures that did not elicit significant spirometric responses have been reported (Holz *et al.*, 2005; McBride *et al.*, 1994)." Agency staff's analyses placed most emphasis on spirometric evidence and not enough emphasis on serious morbidity (*e.g.*, hospital admissions) and mortality observed in epidemiology studies (see page 6-44).

*Therefore, on the basis of the large amount of recent data evaluating adverse health effects at levels at and below the current NAAQS for ozone, it is the unanimous opinion of the CASAC that the current primary ozone NAAQS is not adequate to protect human health. Furthermore, the Ozone Panel is in complete agreement both that: the EPA staff conclusion in Section 6.3.6 arguing that "consideration could be given to retaining the current 8-hr ozone standard" is not supported by the relevant scientific data; and that the current primary 8-hr*

*standard of 0.08 ppm needs to be substantially reduced to be protective of human health, particularly in sensitive subpopulations.*

Additionally, we note that the understanding of the associated science has progressed to the point that *there is no longer significant scientific uncertainty regarding the CASAC's conclusion that the current 8-hr primary NAAQS must be lowered.* A large body of data clearly demonstrates adverse human health effects at the current level of the 8-hr primary ozone standard. Retaining this standard would continue to put large numbers of individuals at risk for respiratory effects and/or significant impact on quality of life including asthma exacerbations, emergency room visits, hospital admissions and mortality. (Scientific uncertainty does exist with regard to the lower level of ozone exposure that would be fully-protective of human health. The Ozone Panel concludes that it is possible that there is no threshold for an ozone-induced impact on human health and that some adverse events may occur at policy-relevant background.)

Moreover, EPA staff concluded that changes in the concentration-based form of the standard (*i.e.*, whether to use the third-, fourth-, or fifth-highest daily maximum 8-hr average concentration) should also be considered. The analysis found in the 2nd Draft Ozone Staff Paper indicates that modest changes in the form of the standard can have substantial impacts on the frequency of adverse health effects. Therefore, the CASAC recommends that the Agency conduct a broader evaluation of alternative concentration-based forms of the primary 8-hr ozone standard and the implications of those alternative forms on public-health protection and stability (*i.e.*, with respect to yearly variability to ensure a stable target for control programs).

The CASAC further recommends that the ozone NAAQS should reflect the capability of current monitoring technology, which allows accurate measurement of ozone concentrations with a precision of parts per *billion*, or equivalently to the third decimal place on the parts-per-million scale. In addition, given that setting a level of the ozone standard to only two decimal places inherently reflects upward or downward "rounding," *e.g.*, 0.07 ppm includes actual measurements from 0.0651 ppm to 0.0749 ppm, the CASAC chooses to express its recommended level, immediately below, to the third decimal place.

*Accordingly, the CASAC unanimously recommends that the current primary ozone NAAQS be revised and that the level that should be considered for the revised standard be from 0.060 to 0.070 ppm, with a range of concentration-based forms from the third- to the fifth-highest daily maximum 8-hr average concentration.* While data exist that adverse health effects may occur at levels lower than 0.060 ppm, these data are less certain and achievable gains in protecting human health can be accomplished through lowering the ozone NAAQS to a level between 0.060 and 0.070 ppm.

## Secondary Ozone NAAQS

An important difference between the effects of acute exposures to ozone on human health and the effects of ozone exposures on welfare is that vegetation effects are more dependent on the *cumulative* exposure to, and uptake of, ozone over the course of the entire growing season (defined to be a minimum of at least three months). *Therefore, there is a clear need for a*

*secondary standard which is distinctly different from the primary standard in averaging time, level and form.* Developing a biologically-relevant ozone air quality index would be directly responsive to the 2004 National Research Council (NRC) recommendations on Air Quality Management in the United States (NAS, 1994) and will help support important new Agency initiatives to enhance ecosystem-related program tracking and accountability.

In its 1996 review of the ozone NAAQS, EPA staff proposed several cumulative seasonal ozone exposure indices, including SUM06, the concentration-weighted metric (*i.e.*, the seasonal sum of all hourly average concentrations > 0.06 ppm), and W126, the integrated exposure index with a sigmoidal weighting function, as candidates for a secondary standard. The Administrator considered a three-month, 12-hr SUM06 secondary standard at a level of 25 ppm-hr as an appropriate, biologically-relevant secondary standard, but ultimately rejected this option in favor of simply setting the secondary standard equal to the primary. It was rationalized that efforts to attain the new 8-hr primary standard would also eliminate most adverse effects on vegetation, and at that time there were uncertainties in how cumulative seasonal exposures would change with efforts to reduce peak 8-hour concentrations. Additionally, it was assumed that future ozone/vegetation effects research over the coming years would clarify the very uncertain quantitative relationships between ozone exposures and vegetation/ecological responses under ambient field conditions.

Unfortunately, however, the Agency has supported very little new vegetation/ecological ozone effects research over the past decade. The net result is that the quantitative evidence linking specific ozone concentrations to specific vegetation/ecological effects must continue to be characterized as having high uncertainties due to the lack of data for verification of those relationships. It is not surprising that substantial research needs remain, as indicated both in Chapter 8 and in individual reviewer comments. The quantitative evidence linking specific ozone concentrations to specific vegetation effects — especially at the complex ecosystem level — must continue to be characterized as having high uncertainties due to the lack of data for verification of those relationships. To a large extent, this is an unavoidable consequence of the inherent complexities of ecosystem structure and function, interactions among biotic and abiotic stressors and stimuli, variability among species and genotype, detoxification and compensatory mechanisms, *etc.* Nevertheless, the compelling weight of evidence provided in Chapter 7 of the 2nd Draft Ozone Staff Paper results from the convergence of results from many various and disparate assessment methods including chamber and free air exposure, crop yield and tree seedling biomass experimental studies, foliar injury data from biomonitoring plots, and modeled mature tree growth.

Despite limited recent research, it has become clear since the last review that adverse effects on a wide range of vegetation including visible foliar injury are to be expected and have been observed in areas that are below the level of the current 8-hour primary and secondary ozone standards. Such effects are observed in areas with seasonal 12-hr SUM06 levels below 25 ppm-hr (the lower end of the range of a SUM06 secondary standard suggested in the 1996 review and the upper end of the range suggested in Chapter 8 of the 2nd Draft Ozone Staff Paper). Seasonal SUM06 (or equivalent W126) ranges well below 25 ppm-hr were recommended for protecting various managed and unmanaged crops and tree seedlings in the 1997 workshop on secondary ozone standards (Heck and Cowling, 1997). The absence of clear-

cut lower effects thresholds for sensitive vegetation combined with the lower recent estimates of policy-relevant background (typical range of 0.015 to 0.035 ppm) emphasizes the importance of efforts to reduce low- to mid-range environmental exposures below 0.060 ppm.

Based on the Ozone Panel's review of Chapters 7 and 8, *the CASAC unanimously agrees that it is* not *appropriate to try to protect vegetation from the substantial, known or anticipated, direct and/or indirect, adverse effects of ambient ozone by continuing to promulgate identical primary and secondary standards for ozone. Moreover, the members of the Committee and a substantial majority of the Ozone Panel agrees with EPA staff conclusions and encourages the Administrator to establish an alternative cumulative secondary standard for ozone and related photochemical oxidants that is distinctly different in averaging time, form and level from the currently existing or potentially revised 8-hour primary standard.* The suggested approach to the secondary standard is a cumulative seasonal growing standard such as the indices SUM06 or W126 aggregated over at least the three summer months exhibiting the highest cumulative ozone levels and includes the ozone exposures from at least 12 daylight hours. The CASAC suggests a range of 10 to 20 ppm-hours for the three-month growing season SUM06 index for agricultural crops rather than the 15-25 ppm-hours proposed in Chapter 8.

However, the Ozone Panel views the three-month growing season W126 index as a potentially more biologically-relevant index than the 3-month growing season SUM06 index. This is because the W126 index has no absolute minimum ozone concentration threshold and only lightly weights the lower ozone concentrations. Therefore, a three-month seasonal W126 that is the approximate equivalent of the SUM06 at 10 to 20 ppm-hr is preferred. As shown by the references cited at the end of Chapter 8, the consensus view among expert persons in the ecological communities of both this country and elsewhere around the world is that *a secondary standard of cumulative form and extending over an entire growing season will be far more effective than a secondary standard that is not cumulative in form and does not include the whole growing season.*

In conclusion, the Clean Air Scientific Advisory Committee is pleased to provide its scientific advice and recommendations to the Agency on the primary and secondary ozone NAAQS. We recognize that our recommendation of lowering of the current primary ozone standard would likely result in a large portion of the U.S. being in non-attainment. *Nevertheless, we take very seriously the statutory mandate in the Clean Air Act not only for the Administrator to establish, but also for the CASAC to recommend to the Administrator, a primary standard that provides for an "adequate margin of safety … requisite to protect the public health."*

Finally, as announced during the Ozone Panel's August meeting, once the Agency releases the Final Ozone Staff Paper in early January 2007, the CASAC intends to hold a public teleconference in late January or early February 2007 for the members of the Ozone Panel to review — and, prospectively, to offer additional, unsolicited advice to the Agency concerning — Chapter 6 (Staff Conclusions on Primary $O_3$ NAAQS) and Chapter 8 (Staff Conclusions on Secondary $O_3$ NAAQS) in that final Agency document. The purpose of such advice would be to

inform EPA's efforts as it develops the forthcoming, proposed rule for ozone and related photochemical oxidants. As always, the CASAC wishes EPA well in this important endeavor.

Sincerely,

/Signed/

Dr. Rogene Henderson, Chair
Clean Air Scientific Advisory Committee

Appendix A – Clean Air Scientific Advisory Committee Roster (FY 2006)

Appendix B – CASAC Ozone Review Panel Roster

Appendix C – Charge to the CASAC Ozone Review Panel

Appendix D – Review Comments from Individual CASAC Ozone Review Panel Members

## CASAC Chapter-Specific Discussion Comments on EPA's 2nd Draft Ozone Staff Paper

Sub-groups of the CASAC Ozone Review Panel who led the discussion on individual chapters of the Staff Paper summarized their comments in the following paragraphs:

**Chapter 2 (Air Quality Characterization):** A better introduction to the central role of photochemical oxidation reactions as the key reactions governing the behavior of air pollutants in the atmosphere would improve this chapter. Ozone is the key indicator of the extent of oxidative chemistry and serves to integrate multiple pollutants. Oxidation in the atmosphere leads to the formation of particulate matter from $SO_2$, $NO_X$, and volatile organic compounds (VOCs) as well as gas phase irritants (formaldehyde, acrolein, etc). Thus, although ozone itself has direct effects on human health and ecosystems, it can also be considered as indicator of the mixture of photochemical oxidants and of the oxidizing potency of the atmosphere. Section 2.2.6 only briefly covers the relationship of ozone to other photochemical oxidants. It would be beneficial to add a short paragraph outlining the role of ozone and other photochemical oxidants in the atmospheric transformation processes that may results in the formation of more toxic products (both in an outdoor and indoor environment), as provided in the individual comments appended to this letter.

The section on policy-relevant background (2.7) continues to have problems. Although the section briefly cites the results of comparison of different models and measurements, it does not adequately address the uncertainties of the global GEOS-CHEM model, and how these uncertainties are reflected in the health risk analysis. Since ozone health effects are observed down to concentrations of the order of 0.04–0.05 ppm, it is important to know how the PRB is related to the considered primary ozone standard and what uncertainties there are in the risk attributed to controllable sources.

**Chapter 3 (Policy-Relevant Assessment of Health Effects Evidence):** The latest draft of Chapter 3 is much improved over the previous draft. Efforts to respond to some of the earlier concerns expressed by the CASAC are appreciated. While in general this chapter is well written, and is a credible basis for the risk analyses that follow, there are inconsistencies and inaccuracies that still need to be addressed. Typically, there is appropriate use of cautionary phrases when the data are not as strong as they might be, but this use is inconsistent across the chapter, and there are instances where EPA staff appear to be stretching to infer that data support their statement. While the individual comments of Ozone Panel members attached to this letter provide specifics on these points, some of the Panel's more significant concerns are discussed briefly below.

Discussion of measurement error is convoluted, confusing, and contains some mistakes. The primary issue in the use of central ambient monitors for ozone in time-series epidemiological studies is whether they provide any information at all that reflects daily personal ozone exposure in susceptible populations. The discussion on p. 3-37 of the impact of various types of exposure measurement error is incorrect; the difference between true and measured ambient concentrations is an example of classical measurement error that results in bias of effect

estimates to the null, not just an increase in standard error. Claiming that the difference between average personal exposure and ambient concentrations results in "attenuation of risk" is not appropriate.

The Ozone Panel does not completely agree with staff's conclusion that "the use of routinely monitored ambient ozone concentrations as a surrogate for personal exposures is not generally expected to change the principal conclusions from ozone epidemiological studies." Indeed, Panel members have little insight as to what we would find if we had actual exposure measurements. Personal exposures most likely correlate better with central site values for those subpopulations that spend a good deal of time outdoors, which coincides, for example, with children actively engaged in outdoor activities, and which happens to be a group that the ozone risk assessment focuses upon.

Some statements about which individuals are at greatest risk of ozone-induced effects are not adequately supported by the information discussed in the chapter. Individuals with chronic obstructive pulmonary disease (COPD) and cardiovascular disease (CVD) are likely to be at increased risk, but the hypothesis that such "hyper-responsiveness" can be used to identify individuals with COPD or CVD who are at greatest risk of $O_3$-induced health effects has not been confirmed. A more appropriate conclusion would be that individuals with COPD and CVD are at increased risk of $O_3$-induced health effects.

The discussion of the ranges for changes in $FEV_1$ that are considered to be small, moderate, or large for persons with impaired respiratory systems is not consistent. While EPA staff state that the table values for the ranges do not need to be changed, staff indirectly acknowledge that a 10% reduction in this variable in asthmatics could have serious consequences, an interpretation that is used in Chapters 4-6.

The 30 subjects studied by Adams had a great influence on the analyses presented in Chapters 5 and 6. While the discussion of the low-level exposures used in the controlled human studies by Adams and colleagues is technically correct that no statistically significant changes were found in $FEV_1$ for ozone at 40 to 60 ppb compared to filtered air, there were clearly a few individuals who experienced declines in lung function at these lower concentrations. These were healthy subjects, so the percentage of asthmatic subjects, if they had been studied, would most likely be considerably greater.

The lack of statistical power is consistently offered in Chapter 3 for why there appears to be an inconsistent effect seen for COPD mortality. Coherence of respiratory effects for ozone suffers from neither no more nor no less power considerations that do those for particulate matter (PM). Yet the Agency did not argue a lack of power when assessing PM risks, so consistency is needed here relative to ozone effect estimates for COPD mortality.

The relatively strong and relatively consistent effect of ozone on emergency department visits for respiratory disease, especially asthma, as evidenced in Figure 3-4 is misrepresented in several places in the Chapter (and in Chapters 5 and 6) as "inconclusive" or "inconsistent." This should be corrected.

**Chapter 4 (Characterization of Human Exposure to Ozone):** The second draft of Chapter 4 has responded to many of the comments made on the first draft, and is thus clearer than before. The panel was pleased to see the reanalysis for 2002 in addition to 2004.

It would be helpful to have the estimated exposures for current (2002 and 2004) levels displayed in Tables 4-8 & 4-9 (p. 4-32) and Figures 4-4 to 4-21 (pp. 4-33 to 4-41), in addition to only those for just meeting the current standard and alternative more stringent standards. This would be analogous to the way estimated effects are displayed in Chapter 5 (Figures 5-5 to 5-9 [pp.5-58 to 5-65]).

On the whole, Chapter 4 provides a clear "road map" for what was done to characterize available knowledge about human exposure to ozone in the framework of generally accepted modeling approaches of appropriately selected populations in 12 urban areas of the U.S. Much of the text reads like a basic textbook on human exposure assessment using state-of-the-art modeling approaches, such as the Air Pollutants Exposure Model (APEX), including adjustments for lung ventilation of delivered ozone dose. This extension, beyond exposure characterization, is particularly important for ozone where the extent of measurable human responses is very sensitive to the amount of ozone inhaled and to where it deposits along the respiratory tract. Further extension of the methodology to estimate dose would have important implications and should be discussed.

There is an explicit discussion of the limitations of the APEX model in terms of variability and the quality of the input data, which is appropriate and fine as far as it goes. There are good reasons presented for selection of urban areas and the time periods to be modeled. However, there was inadequate consideration of the populations selected for modeling. Those selected were appropriate, but the omission of the elderly, the population most at risk for ozone-associated premature daily mortality, was notable and not even mentioned in terms of why it was not considered.

The chapter was very good at exposition and clear presentation of modeling results, but was deficient it its discussion of seemingly counterintuitive results, and of a potentially large influence of measurement biases. As an example of the first of these issues, the children in LA & Houston are estimated to have far fewer exposures above 0.07 ppm (8-hr) than in most other cities with lower ozone concentrations and fewer children. This was likely due to the greater within-day and sampler-to-sampler variations in concentration within these two cities than in the others, the fact that the entire year was modeled while for other sites the winter was not included and/or the greater extent of air conditioning, especially in Houston. Whatever the reasons, there should have been some discussion of the causes. The quadratic rollback methodology should have been better described since this strategy has important consequences for the modeled results.

The second issue that was presented, but left hanging without an adequate discussion is at the bottom of page 4-47, where it was simply stated that "in general, APEX systematically under-predicts the measured values by 0.001 to 0.02 ppm (zero to 50 percent)." If this is so, is it due to a really serious failure of the APEX model, or to unreliable measurements? The measurements at issue were six-day average concentrations based on the use of passive

(diffusion) samplers, which are known to be subject to significant errors when the air velocity across the inlet is variable. The comparison of measured and modeled concentrations depicted in Figure 4-22 is certainly worthy of further analysis and discussion.

**Chapter 5 (Characterization of Health Risks):** Generally the panel found Chapter 5 and its accompanying risk assessment to be well done, balanced and reasonably communicated. Additional text is needed at the beginning and end of the chapter to put the limited risk assessment into the context of the much larger body of evidence of ozone health effects. The discussion of uncertainty in these risk estimates is expanded in section 5.3.2.5. Although a number of issues are raised, their impacts on the estimates have not been thoroughly explored. Additional sensitivity analyses seem warranted. In particular, it is essential that the sensitivity of the risk assessment to the shape of the dose-response curve for $FEV_1$ be evaluated. Although the 3 parameter logistic (3PL) model emulates the pattern seen in the five "data points," these points are aggregates of the original data, and may give a misleadingly optimistic picture of the quality of the fit. More importantly, although the problem of model uncertainty is noted it has not been addressed even though methods exist for doing so. Even if only the linear and logistic models were included in the analysis, the error bands around the estimated response probabilities would likely increase to better reflect that uncertainty. In addition, a suggestion to deal with the uncertainties surrounding estimation of PRB, particularly as related to Table 5.5 (for lung function) and Table 5.11 (mortality), would be to change the form of the analyses to assess the impact of the concentration change in the expected number of health effects relative to the current standard. The key advantage of estimating the effect of concentration change is that it does not depend on the choice of the PRB.

With regard to the controlled human exposure studies, Ozone Panel members believe that the selection of changes in pulmonary function expressed as percent change in $FEV_1$ in children is a fair indicator of an adverse effect at 15% change in all active children; and, in asthmatic children, a 10% change is indicative of adverse effects. However, the presentation of the figures showing these effects needs to be revised to indicate the uncertainties in the results used, particularly at the lower levels of exposure. The potential mechanisms whereby these changes are a reflection of both pain on breathing, partial inflammation of smaller airways, other effects on airways, and potentially triggers for more significant respiratory morbidity, particularly in asthmatic children, are not adequately discussed. In addition, some added discussion is necessary to indicate that these measures are generally taken in areas with relatively high background levels of ozone exposure, and that the role that tolerance may play in minimizing the degree of adverse effect observed needs to be considered.

From the perspective of the epidemiological data, the Ozone Panel judged the selection of: respiratory symptoms in moderate/severe asthmatic children (ages zero [birth] to 12); hospital admissions for respiratory illness among asthmatic children; and premature total non-accidental and cardiorespiratory mortality for inclusion in the quantitative risk assessment to be appropriate. However, the CASAC believes that several other endpoints should be discussed qualitatively to support the findings that these endpoints indicate that significant adverse effects are occurring at exposure concentrations well below the current standard. Other endpoints deemed worthy of additional discussion included respiratory emergency department visits among asthmatics and patients with other respiratory diseases, increased medication usage, and increased

symptomatology reported at exposure levels well below the current standard. Taken together, members of the Ozone Panel felt strongly that these findings preclude including the current standard as a scientifically defensible option for the Administrator (see discussion about Chapter 6 found in the main portion of the letter above).

Another problem in the health effects calculations (see Table 5-5 and 5-11) is that they are based on computations of the form $R_x - R_{PRB}$, where $R_x$ is the risk at a given concentration x of $O_3$ and $R_{PRB}$ is the corresponding risk at policy-relevant background (PRB) for $O_3$. As discussed at the Ozone Panel's August meeting, the PRB is highly-problematic to calculate and is, in some sense, "unknowable." One can avoid this problem by calculating the $\Delta = R_{0.8} - R_x$ for various concentrations x. This form would allow focus on the question, "What is the difference in the expected number of health effects that will occur at various concentrations of $O_3$, relative to the current standard of 0.08?" A key advantage of $\Delta$ is that it does not depend on the choice of PRB, and thus is free of the uncertainties surrounding estimation of PRB.

**Chapter 6 (Staff Conclusions on Primary $O_3$ NAAQS):** See the discussion on Chapter 6 found in the main portion of the letter above. It would also be helpful to have the estimated exposures for current (2002 and 2004) levels displayed in figures 6-1 to 6-6 (pp. 6-34 to 6-39), in addition to only those for just meeting the current standard and alternative more stringent standards. This would be analogous to the way estimated effects are displayed in Chapter 5 (Figures 5-5 to 5-9 [pp.5-58 to 5-65]).

**Chapters 7 (Policy-Relevant Assessment of Welfare Effects Evidence) and 8 (Staff Conclusions on Secondary $O_3$ NAAQS):** Chapter 7 is a well-developed and persuasively presented assessment of the welfare effects of ozone on vegetation, which forms the technical basis for the range of secondary standards recommended in Chapter 8. That having been said, the potential for significant propagation of error/uncertainty in the underlying technical documentation draws into question the conclusions drawn by EPA Staff. As observed in the Agency's 1989 and 1996 Ozone Staff Papers, ozone remains the most prevalent phytotoxic compound in the ambient air "impairing crop production and injuring native vegetation and ecosystems more than any other air pollutant" (USEPA 1989, 1996). Furthermore, as has been noted in the current assessment of human health effects, there also appears to be no safe threshold concentration below which ozone effects on sensitive vegetation are eliminated. See the additional discussion on Chapter 8 found in the main portion of the letter above.

# Appendix A – Clean Air Scientific Advisory Committee Roster (FY 2006)

## U.S. Environmental Protection Agency
## Science Advisory Board (SAB) Staff Office
## Clean Air Scientific Advisory Committee (CASAC)

### CHAIR

**Dr. Rogene Henderson**, Scientist Emeritus, Lovelace Respiratory Research Institute, Albuquerque, NM

### MEMBERS

**Dr. Ellis Cowling**, University Distinguished Professor-at-Large, North Carolina State University, Colleges of Natural Resources and Agriculture and Life Sciences, North Carolina State University, Raleigh, NC

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

**Dr. Frederick J. Miller**, Consultant, Cary, NC

**Mr. Richard L. Poirot**, Environmental Analyst, Air Pollution Control Division, Department of Environmental Conservation, Vermont Agency of Natural Resources, Waterbury, VT

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

**Dr. Barbara Zielinska**, Research Professor, Division of Atmospheric Science, Desert Research Institute, Reno, NV

### SCIENCE ADVISORY BOARD STAFF

**Mr. Fred Butterfield**, CASAC Designated Federal Officer, 1200 Pennsylvania Avenue, N.W., Washington, DC, 20460, Phone: 202-343-9994, Fax: 202-233-0643 (butterfield.fred@epa.gov) (Physical/Courier/FedEx Address: Fred A. Butterfield, III, EPA Science Advisory Board Staff Office (Mail Code 1400F), Woodies Building, 1025 F Street, N.W., Room 3604, Washington, DC 20004, Telephone: 202-343-9994)

# Appendix B – CASAC Ozone Review Panel Roster

## U.S. Environmental Protection Agency
## Science Advisory Board (SAB) Staff Office
## Clean Air Scientific Advisory Committee (CASAC)
## CASAC Ozone Review Panel

**CHAIR**

**Dr. Rogene Henderson***, Scientist Emeritus, Lovelace Respiratory Research Institute, Albuquerque, NM

**MEMBERS**

**Dr. John Balmes**, Professor, Department of Medicine, University of California San Francisco, University of California – San Francisco, San Francisco, California

**Dr. Ellis Cowling***, University Distinguished Professor-at-Large, North Carolina State University, Colleges of Natural Resources and Agriculture and Life Sciences, North Carolina State University, Raleigh, NC

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

**Dr. William (Jim) Gauderman**, Associate Professor, Preventive Medicine, Medicine, University of Southern California, Los Angeles, CA

**Dr. Henry Gong**, Professor of Medicine and Preventive Medicine, Medicine and Preventive Medicine, Keck School of Medicine, University of Southern California, Downey, CA

**Dr. Paul J. Hanson**, Senior Research and Development Scientist, Environmental Sciences Division, Oak Ridge National Laboratory (ORNL), Oak Ridge, TN

**Dr. Jack Harkema**, Professor, Department of Pathobiology, College of Veterinary Medicine, Michigan State University, East Lansing, MI

**Dr. Philip Hopke**, Bayard D. Clarkson Distinguished Professor, Department of Chemical Engineering, Clarkson University, Potsdam, NY

**Dr. Michael T. Kleinman**, Professor, Department of Community & Environmental Medicine, University of California – Irvine, Irvine, CA

**Dr. Allan Legge**, President, Biosphere Solutions, Calgary, Alberta, Canada

**Dr. Morton Lippmann**, Professor, Nelson Institute of Environmental Medicine, New York University School of Medicine, Tuxedo, NY

**Dr. Frederick J. Miller***, Consultant, Cary, NC

**Dr. Maria Morandi**, Assistant Professor of Environmental Science & Occupational Health, Department of Environmental Sciences, School of Public Health, University of Texas – Houston Health Science Center, Houston, TX

**Dr. Charles Plopper**, Professor, Department of Anatomy, Physiology and Cell Biology, School of Veterinary Medicine, University of California – Davis, Davis, California

**Mr. Richard L. Poirot***, Environmental Analyst, Air Pollution Control Division, Department of Environmental Conservation, Vermont Agency of Natural Resources, Waterbury, VT

**Dr. Armistead (Ted) Russell**, Georgia Power Distinguished Professor of Environmental Engineering, Environmental Engineering Group, School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA

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

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

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

**Dr. Sverre Vedal**, Professor of Medicine, Department of Environmental and Occupational Health Sciences, School of Public Health and Community Medicine, University of Washington, Seattle, WA

**Dr. James (Jim) Zidek**, Professor, Statistics, Science, University of British Columbia, Vancouver, BC, Canada

**Dr. Barbara Zielinska***, Research Professor, Division of Atmospheric Science, Desert Research Institute, Reno, NV

## SCIENCE ADVISORY BOARD STAFF

**Mr. Fred Butterfield**, CASAC Designated Federal Officer, 1200 Pennsylvania Avenue, N.W., Washington, DC, 20460, Phone: 202-343-9994, Fax: 202-233-0643 (butterfield.fred@epa.gov)

* Members of the statutory Clean Air Scientific Advisory Committee (CASAC) appointed by the EPA Administrator (FY 2006)

# Appendix C – Charge to the CASAC Ozone Review Panel

---

$O_3$ air quality information and analyses (Chapter 2):

1. To what extent are the air quality characterizations and analyses clearly communicated, appropriately characterized, and relevant to the review of the primary and secondary $O_3$ NAAQS?

2. Does the information in Chapter 2 provide a sufficient air quality-related basis for the exposure, human health and environmental effects, health risk assessment, and environmental assessment presented in later chapters?

$O_3$-related health effects (Chapter 3):

1. To what extent is the presentation of evidence from the health studies assessed in the AQCD and the integration of information from across the various health-related research areas drawn from the $O_3$ AQCD technically sound, appropriately balanced, and clearly communicated?

2. What are the views of the Panel on the appropriateness of staff's discussion and conclusions in Chapter 3 on key issues related to quantitative interpretation of animal toxicology and controlled-exposure human experimental studies and epidemiologic study results, including, for example, exposure error, the influence of alternative model specification, potential confounding or effect modification by co-pollutants, and lag structure?

3. What are the Panel's view on the adequacy and clarity of staff discussion on the issue of potential thresholds in concentration-response relationships discussed in Chapter 3?

Exposure Analysis (Second Draft Chapter 4 of the $O_3$ Staff Paper, draft Exposure Analysis technical support document, and OAQPS Staff Memorandum on Uncertainty Analysis):

1. To what extent are the assessment, interpretation, and presentation of the results of the exposure analysis as presented in Chapter 4 (and in the second draft Exposure Analysis technical support document) technically sound, appropriately balanced, and clearly communicated?

2. Are the methods used to conduct the exposure analysis technically sound? Does the Panel have any comments on the methods used?

3. To what extent are the uncertainties associated with the exposure analysis clearly and appropriately characterized in Chapter 4, the Exposure Analysis technical support document, and the uncertainty memorandum?

4. To what extent is the plan for the remaining uncertainty assessment technically sound? Are there other important uncertainties which are not covered? What are the views of the Panel on sensitivity analyses conducted to evaluate the influence of uncertainties in the exposure analysis?

Heath Risk Assessment (Second Draft Chapter 5 of the $O_3$ Staff Paper and draft Health Risk Assessment technical support document):

1. To what extent are the assessment, interpretation, and presentation of the results of the revised exposure analysis as presented in Chapter 5 (and in the second draft Risk Assessment technical support document) technically sound, appropriately balanced, and clearly communicated?

2. In general, is the set of health endpoints and concentration-response and exposure-response functions used in this risk assessment appropriate for this review?

3. Are the methods used to conduct the health risk assessment technically sound? Does the Panel have any comments on the methods used?

4. To what extent are the uncertainties associated with the health risk assessment clearly and appropriately characterized in both the second draft Chapter 5 and the second draft Health Risk Assessment technical support documents?

Staff Conclusions and Standard Options for the Primary $O_3$ NAAQS (Chapter 6):

1. What are the views of the Panel on the approach taken by staff (as discussed in Chapter 6) of using both evidence-based and quantitative exposure- and risk-based considerations in drawing conclusions and identifying options as to a range of standards to protect against health effects associated with exposure to $O_3$, alone and in combination with the ambient mix of photochemical oxidants, for consideration in this review of the primary $O_3$ NAAQS?

2. Does the Panel generally agree that the range of alternative primary $O_3$ standards identified in Chapter 6 is generally consistent with the available scientific information and is appropriate for consideration by the Administrator?

3. What are the views of the Panel on the key uncertainties and $O_3$ research recommendations discussed in Chapter 6?

$O_3$-related welfare effects and secondary standard options (Chapters 7):

1. To what extent is the presentation of evidence drawn from the vegetation effects studies assessed in the $O_3$ AQCD technically sound, appropriately balanced, and clearly communicated?

2. What are the views of the Panel on the appropriateness of staff's weight-of-evidence approach which assesses information from across the various vegetation-related research areas described in the $O_3$ AQCD, including chamber and free air exposure crop yield and tree seedling biomass experimental studies, foliar injury data from biomonitoring plots, and modeled mature tree growth?

3. To what extent are the methods used to conduct the exposure assessment and the interpretation and presentation of the results of the exposure assessment in the second draft Chapter 7 and the draft Environmental Assessment technical support document technically sound, appropriately balanced, and clearly communicated?

4. To what extent are the uncertainties associated with the exposure analysis clearly and appropriately characterized in the second draft Chapter 7 and the draft Environmental Assessment technical support document?

5. To what extent are the uncertainties associated with the vegetation risk assessment clearly and appropriately characterized in both the second draft Chapter 7 and the draft Environmental Assessment technical support document?

6. Staff recognizes that gradients can exist between $O_3$ levels measured at monitor probe heights and those measured over low vegetation canopies. What are the Panel's views on the appropriateness of applying a single adjustment factor to hourly monitoring data to account for the range of potential gradients that can exist across sites and crop and tree seedling canopy structures? Are there alternative approaches or adjustment values the Panel would suggest? Are staff's planned sensitivity analyses appropriate and sufficient?

7. To what extent do the figures aid in clarifying the text? Should more or less information of this type be included in the final Chapter 7 or its Appendices?

8. Given the lack of quantitative information on $O_3$-related ecosystem effects, what are the Panel's views on the appropriateness of how this topic is addressed in the second draft Chapter 7?

<u>Staff Conclusions and Standard Options for the Secondary $O_3$ NAAQS (Chapter 8)</u>:

1. Does the Panel generally agree that the secondary standard options identified by staff (including indicator, averaging time, form, and level) are generally consistent with the available scientific and technical information and are appropriate for consideration by the Administrator?

# Appendix D – Review Comments from Individual CASAC Ozone Review Panel Members

This appendix contains the preliminary and/or final written review comments of the individual members of the Clean Air Scientific Advisory Committee (CASAC) Ozone Review Panel on the 2nd Draft Ozone Staff Paper and three related draft technical support documents — *Ozone Health Risk Assessment for Selected Urban Areas: Draft Report* (2nd Draft Ozone Health Risk Assessment, July 2006); *Ozone Population Exposure Analysis for Selected Urban Areas: Draft Report* (2nd Draft Ozone Exposure Assessment, July 2006); and *Draft Ozone Environmental Assessment: Exposure, Risk and Benefits Assessment* (Draft Ozone Environmental Assessment, July 2006) — who submitted such comments electronically. The comments are included here to provide both a full perspective and a range of individual views expressed by Panel members during the review process. These comments do not represent the views of the CASAC Ozone Review Panel, the CASAC, the EPA Science Advisory Board, or the EPA itself. The views of the CASAC Ozone Review Panel and the CASAC as a whole are contained in the text of the report to which this appendix is attached. Panelists providing review comments are listed on the next page, and their individual comments follow.

**Panelist** **Page #**

# Dr. John Balmes

Review of Chapter 3 of the 2nd draft Ozone staff paper – John Balmes

Overall, I find this draft of Chapter 3 much improved over the previous draft. In particular, the efforts to respond to the some of the concerns expressed in the letter written by CASAC subsequent to the publication of the CD are appreciated.

In general, the chapter provides careful interpretation of the available data in the controlled human exposure, epidemiological, and animal toxicological literature. However, I do think that several of the statements about which individuals are at greatest risk of ozone-induced effects in the concluding paragraph of the chapter are not adequately supported by the information discussed in the chapter. First, we do not know that individuals "with cardiorespiratory impairment (e.g., those with COPD or cardiovascular disease) who are "hyperresponsive" to $O_3$ exposure (i.e., exhibit much higher than normal lung function decrements and/or respiratory symptoms) would be considered at greatest risk of $O_3$ exposure." While I agree that individuals with COPD and cardiovascular disease (CVD) are likely to be at increased risk, I do not believe that the hypothesis that such "hyperresponsiveness" can be used to identify individuals with COPD or CVD who are at greatest risk of $O_3$-induced health effects has been confirmed. The hypothesis is certainly plausible, but for the purpose of this chapter, I suggest that it simply be stated in the conclusion that individuals with cardiorespiratory impairment (e.g., those with COPD or CVD) are at increased risk of adverse effects of $O_3$ exposure.

Second, the final sentence of the chapter states that "those with genetic polymorphisms for antioxidant enzymes and inflammatory genes may be at heightened risk of effects of $O_3$." While I concur with this statement, the chapter provides no supporting evidence from human studies (either controlled exposure or epidemiological). The only relevant supporting information that is discussed in the previous sections of the chapter is that murine studies have used strain differences in $O_3$-induced responses to identify candidate genes that are likely to be involved in determining susceptibility. These studies have involved genes in the early inflammatory response rather than antioxidant enzymes. The previous discussion also includes the caveat that the relevant murine studies used relatively high-dose exposure protocols. Unless, the limited human data suggesting that polymorphisms of certain antioxidant enzyme genes (e.g., GSTM1 and NQO1) are discussed in the chapter, I would revise the final sentence.

Specific Comments

3-6-21 and 22 The Adams et al. studies were not done in southern CA. They were done at UC Davis.

3-9-1 I would add "short-term" before declines in lung function in this sentence to be absolutely clear.

3-9-4 and 5 For both clarity and consistency, I would revise this sentence as follows: "…pulmonary function *in asthma*, which included increased *nonspecific airway responsiveness* secondary to airway inflammation due to $O_3$ exposure."

3-12-7 and 8 This sentence is misleading as currently written. It should be revised as follows: "Both human and animal studies indicate that *AHR is* not associated with airway inflammation…"

3-34-2 This sentence overstates the limitations of the data on ED visits for respiratory disease, especially asthma. For example, in Figure 3-4 ED visits for asthma, only one study did not show an increased risk. I would revise the sentence as follows: "…positive *but somewhat less robust* evidence for associations with respiratory ED visits."

3-36-9 and 10 This sentence is a bit unclear as written. I would revise this sentence to end after "…studies."

3-48-11 Spirometry cannot be done in animals so "measures of lung function" should be substituted.

3-56-33 Controlled human exposure studies cannot evaluate effects of long-term $O_3$ exposures.

3-57-15 through 17 In my view, this sentence misrepresents the status of the available evidence. I would favor a statement that used "inconclusive" or "limited" to describe the evidence on long-term $O_3$-induced effects on lung function.

3-60-35 Change "don't" to *do not*.

3-60-32 through 34 This sentence is poorly written and should be revised as follows: "Controlled studies of dietary antioxidant supplements have shown some protective effects on lung function decrements but not on symptoms and airway inflammatory responses."

3-64-16 through 18 The Vagaggini study cited exposed subjects to allergen before $O_3$ and obtained samples of respiratory tract lining fluid by sputum induction rather than by bronchoalveolar lavage.

3-65-3 This sentence is simply untrue regarding asthmatic subjects. I would revise the sentence to delete the words "and asthmatic".

3-67-11 This sentence is poorly written. I would revise the sentence as follows: "…the involvement *of specific genes or genetic loci* in $O_3$-induced airway hyperresponsiveness and inflammation."

Table 3-3 For consistency, I would replace "bronchial" with "airway".

3-75-3 through 5 These bullets should be revised to "…*$O_3$ after...* "

3-81-16 Use "*ours*" here rather than "our."

3-82-27 through 29 This statement is again too strong in my view. I would suggest wording to the effect that the data from epidemiological studies is "limited" or "inconclusive" while data from clinical studies is "lacking".

3-82-30 and 31 This sentence is poorly written. I would revise the sentence as follows: "…of $O_3$*-induced* effects with a clean air *recovery period*…"

References The Mortimer reference listed is a secondary paper for the study. In addition to the 2002 *Eur Respir J* citation, the primary paper for this study published in 2000 in the *Am J Respir Crit Care Med* should also be cited.

# Dr. Ellis Cowling

Dr. Ellis Cowling
North Carolina State University
August 17, 2006

**General Comments on Chapters 7 and 8 of the "Second Draft OAQPS Staff Paper on Ozone: Policy Assessment of Scientific and Technical Information" and the draft technical support document titled "Technical Report on Ozone Exposure, Risk, and Impacts Assessments for Vegetation: Draft Report"**

The following General and Specific comments are focused almost entirely on the CASAC "Charge Questions" raised in Karen Martin's letter to Fred Butterfield dated July 17, 2006. Please note in the paragraphs below, that Karen Martin's questions are written in **bold-faced non-indented paragraphs**, whereas my responses are written in indented paragraphs without bold type.

**Comments in Response to the Charge Questions for Chapter 7 in the Second Draft Staff Paper on Ozone and the draft technical support document mentioned above:**

**1. To what extent is the presentation of evidence drawn from the vegetation effects studies assessed in the ozone Criteria Document technically sound, appropriately balanced, and clearly communicated?**

I find the evidence drawn from the vegetation effects text and figures of Chapter 7 to be technically sound, appropriately balanced, and clearly communicated.

**2. What are the views of the Panel on the appropriateness of staff's weight-of-evidence approach which assesses information from across various vegetation-related research areas described in the ozone Criteria Document, including chamber and free air exposure crop yield and tree seedling biomass experimental studies, foliar injury data from bio-monitoring plots , and modeled mature tree growth?**

I find the weight-of evidence approach used by EPA staff appropriate and reasonably thorough with regard to chamber and free air exposure crop yield and tree seedling biomass experimental studies, foliar injury data from bio-monitoring plots, and modeled mature tree growth.

**3. To what extent are the methods used to conduct the exposure assessment and the interpretation and presentation of results of the exposure assessment in the second draft Chapter 7 and the draft Environmental Assessment technical support document technically sound, appropriately balanced, and clearly communicated?**

I find these methods, their interpretation, and presentation of results to be technically sound, appropriately balanced, and clearly communicated in both Chapter 7 and the technical support document mentioned above.

**4. To what extent are the uncertainties associated with the exposure analysis clearly and appropriately characterized in both the second draft Chapter 7 and the draft Environmental Assessment technical support document?**

I find these uncertainties to be both clearly and appropriately characterized in both the second draft Chapter 7 and the draft Environmental Assessment technical support document.

**5. To what extent are the uncertainties associated with the vegetation risk assessment clearly and appropriately characterized in both the second draft Chapter 7 and the draft Environmental Assessment technical support document?**

Similarly, I find these additional uncertainties with regard to the vegetation risk assessment to be both clearly and appropriately characterized in both the second draft Chapter 7 and the draft Environmental Assessment technical support document.

**6. Staff recognizes that gradients can exist between ozone levels measured at monitoring probe heights and those measured over low vegetation canopies. What are the Panel's views on the appropriateness of applying a single adjustment factor to hourly monitoring data to account for the range of potential gradients that can exist across sites and crop and tree seedling canopy structures? Are there alternative approaches or adjustment values that the Panel would suggest? Are staff sensitivity analyses appropriate and sufficient?**

I agree that these gradients in ambient air concentrations within vegetation canopies can and do exist and also find that the idea of applying a single adjustment factor to account for these gradients is satisfactory. I think the approach used in these documents is as reasonable as any other that could be considered and that the sensitivity analyses used by staff are both appropriate and sufficient.

**7. To what extent do the figures aid in clarifying the text? Should more or less information of this type be included in the final Chapter 7 or its appendices?**

As indicated in my General Comments on the First Draft Staff Paper on Ozone on December 17, 2005, lack of figures -- and especially pictures and charts displaying the important and sometimes very severe impacts of ozone on crops, forest and shade trees, and natural ecosystems -- is a major shortcoming of Chapter 7 in this Second Draft Staff Paper on Ozone effects on vegetation. Pictures are indeed worth a thousand words (or more!)

I assert once again that it is very important that the final draft of Chapter 7 in this Staff Paper on Ozone should include illustrations that show:

a) The distribution of stomata on leaf surfaces, ideally also a paired set showing stomata open and stomata closed, and a diagram showing the structure of the stomata within the palisade layer of a leaf,

b) Crop plants or floral plants grown under different concentrations of ozone,

c) Differences in grain yield with and without exposure to injurious concentrations of ozone,

d) The greater effect of ozone exposure on root growth than on shoot growth,

c) A collage of pictures showing foliar injury that decreases economic value of various leafy vegetables like spinach, cut flowers, and Christmas trees, and most important of all:

1) Widespread ozone-induced mortality or ozone-and-biotic-pathogen induced mortality of different species trees exposed to ozone in some of the national parks, state parks, and/or wilderness areas in various parts of the US.

2) A picture of the cross-section of the stem of a tree showing differences in width of annual rings in trees grown with and without exposure to injurious concentrations of ozone, and, if possible, also, a cross-section showing the differential width of annual rings on the same tree during years when ozone exposures were high and ozone exposures were low.

**8. Given the lack of quantitative Information on ozone-related ecosystem effects, what are the Panel's views on the appropriateness of how this topic is addressed in the second draft Chapter 7?**

I think the staff have done about as well as can be expected with this question under the present circumstances of substantial scientific and technical ignorance about whole-ecosystem effects.

**Comments in Response to the single Charge Question for Chapter 8 in the the Second Draft Staff Paper on Ozone:**

**1. Does the Panel generally agree that the secondary standard options identified by staff (including indicator, averaging time, form, and level) are generally consistent with the available scientific and technical information and are appropriate for consideration by the Administrator?**

I believe that EPA staff have done precisely what should be done with regard to recommending firmly and persuasively to the Administrator of EPA that the time has come to formulate and implement a Secondary (Welfare-Based) National Ambient Air Quality Standard for Ozone that is distinct in averaging time, form, and level from the Primary Standard.

The scientific and technical evidence in support of the staff recommendations for serious consideration of the alternative cumulative SUM06 and/or W126 Secondary Standards is well-developed and persuasively presented.

As shown by the references cited at the end of Chapter 8, the consensus view among expert persons in the ecological communities of this country and elsewhere around the world is that a Secondary Standard of cumulative form and extending over a whole growing season -- will be far more effective than a Secondary Standard that is not cumulative in form and does not include the whole growing season. Thus, it is simply NOT appropriate to continue to try to protect

vegetation from the substantial, known or anticipated, direct and/or indirect, adverse effects of ozone by continuing to promulgate identical Primary and Secondary Standards for ozone.

This generalization is true for a wide variety of commercially important crop plants. But it also is true for the vegetation in:

a) both intensively managed and wild-land forests,
b) scenic vistas in natural parks and wilderness areas,
c) ornamental shrubs and shade trees in urban, suburban, and rural areas, as well as
d) the vegetation in natural grasslands, rangelands, and other natural ecosystems

all over the United States.

**A few final comments on the relationship between Chapter 2 and Chapters 7 and 8 in this Second Draft Staff Paper on Ozone and the Name of the Standard to which this Second Draft Staff Paper is addressed:**

**1. Relationship between Chapter 2 titled "*Air Quality Characterization*" and Chapter 7 titled *Policy Relevant Assessment of Welfare Effects Evidence* and Chapter 8 titled *Staff Conclusions on Secondary Ozone NAAQS*.**

There are several important parts of the *Air Quality Characterization* (Chapter 2) of this Second Draft Staff Paper on Ozone that are highly relevant to the conclusions drawn in Chapters 7 and 8. Thus, in my General Comments on the First Draft Staff Paper on Ozone, I recommended that maps be included in either Chapter 2 or Chapter 7 (or both) (and possibly also in Chapter 8) that would show the distribution of counties in various parts of the US where exceedances of various alternative Primary and Secondary standards have occurred in the past and/or would be expected to occur in the future.

I had in mind that these maps would show the distribution of counties that exceed: a) the present identical primary and secondary ozone standards (8-hour) and b) and c) -- the two distinctive cumulative secondary standards that were proposed for consideration by the Administrator of EPA in 1996 and again in this Second Draft Staff Paper in 2006 [b) SUM06 and c) W126].

At present, page 7-15 in Chapter 7 is the only place in this Second Draft Staff Paper where I could find all three of these alternative existing and proposed ozone standards defined in way that allows ready comparison among these three alternatives:

a) the existing identical Primary and Secondary Standards,

b) the proposed alternative cumulative SUM06 Secondary Standard, and

c) the proposed alternative cumulative W126 Secondary Standard.

But these definitions are now presented in Chapter 7 – which is a very long way away from the maps that are presented in Chapter 2!

A map showing the distribution of counties across the US where exceedances of the existing identical Primary and Secondary Standards were exceeded in 2002-2004 is shown in Figure 2-6 on page 2-20 in Chapter 2. And several different maps (based on two different sets of

monitoring sites in 2001 and 2002) are shown in Chapter 2 for the SUM06 proposed cumulative Secondary Standard.

But no maps at all are shown in either Chapter 2 or chapters 7 or 8 for the W126 alternative proposed cumulative Secondary Standard. I know (and maybe some readers of the Final Staff Paper on Ozone will also know) that the SUM06 and W126 maps will be much the same, but that very similarity is good reason for showing how very similar they are (or by chance are not so similar in one or another part of the US).

I believe the value of various parts of Chapter 2 for understanding the major conclusions in Chapters 7 and 8 will be enhanced if more inter-chapter linkages and cross-references regarding exposures of crops, forests and natural ecosystems are included in the text of both chapters.

I believe it also will be valuable to include in Chapter 2 or Chapter 7, a chart similar to those on pages 2-28 and 2-32 in the First Draft Staff Paper on Ozone showing the diurnal patterns of ozone concentrations at a high-elevation forest site where ozone concentrations during nighttime hours frequently remain relatively high. These nighttime concentrations of ozone are important because of their injurious effects on high-elevation forests as discussed on page 7-22 of the First Draft Staff Paper on Ozone.

**2) Title of this Draft Staff Paper on Ozone**

In my General Comments on the First Staff Paper on December 17, 2005, I agreed with Philip Hopke's assertion that the title of this Staff Paper should be changed so it will be faithful to the title of the National Ambient Air Quality Standards to which this Staff Paper applies. Since the Standards dealt with in this Staff Paper on Ozone bear the name "Ozone and Related Photochemical Oxidants," both Philip Hopke and I advised that the title of the staff paper should also include "Ozone and Related Photochemical Oxidants."

Apparently, either both Philip and I were wrong in believing that "Related Photochemical Oxidants" are a part of the present name of these NAAQS, or the authors of the Second Draft Staff Paper considered our recommendation not to be relevant to the job they had at hand.

Section 1.2.2 is titled "History of Ozone NAAQS Reviews." But it does not explain when the NAME of these Standards (as opposed to the specific "indicator" for these standards) was changed. Thus, I recommend that Section 1.2.2 be revised to avoid the confusion that at least Philip Hopke and I have had (and continue to have) about the NAME of the standards to which this Staff Paper is relevant.

# Dr. James Crapo

**Comments on Second Draft of the Ozone Staff Paper**
**Chapter 6, "Staff Conclusions on Primary Ozone NAAQS"**

Dr. James Crapo, CASAC
August 22, 2006

This chapter is in general well written and describes the substantial progress that has occurred in the science underlying our understanding of potential health effects from exposures to ozone at levels near or below the current primary standard of 0.08 PPM on an 8 hour average. In the interval since the current standard was set, a variety of studies have found that measurable lung function and/or other health effects occur, particularly in sensitive populations, for exposures to ozone that are at or below the current standard and possibly as low as 0.04 PPM. Uncertainties increase as the level of ozone exposure decreases and the staff are appropriate in considering this in making final recommendations.

The summary recommendations in Section 6.3.6 provide a rationale for both retaining the current 8 hour ozone standard and for decreasing the standard to a level of 0.07 PPM with a range of forms from the third to the fifth highest daily maximum 8 hour average concentration. Given the consistency of data from multiple sources demonstrating adverse health effects at levels at and below the current 0.08 PPM 8 hour standard, the reasonableness of retaining the current standard should be re-assessed. In my opinion, this position is not well justified by the CD or by the arguments in the staff paper. Clear health benefits are estimated by the EPA staff analysis to occur if the standard were lowered to 0.07 PPM and, although uncertainty increases, further benefit is demonstrated to occur by the staff analysis if the standard is lowered to 0.06 PPM. Based on staff analysis that enhanced benefit occurs by lowering the standard as low as 0.06 PPM, I would recommend that the range of consideration for modification of the standard be expanded to include a full analysis of the potential benefit of lowering the standard to 0.06 PPM.

The staff paper should also further consider the form of the standard. The staff paper analysis indicates that modest changes in the form of the standard can have profound impacts on the frequency of adverse health events. This would justify a more thorough evaluation of the form of the standard and a possible recommended change. This is an area where research recommendations should emphasize the uncertainties and opportunity to more rigorously test the local and regional impacts of different choices for the form of the primary ozone standard.

Overall, Section 6 is excellent and provides a strong rational basis for rule setting with regard to the ozone NAAQS.

## Dr. William (Jim) Gauderman

Comments on Staff Paper, post meeting
Gauderman
8/25/06

The health-effects calculations given for lung function (e.g. Table 5-5) and mortality (e.g. Table 5-11) are based on computations of the form $R_x - R_{PRB}$ where $R_x$ is the risk at a given concentration x of $O_3$ and $R_{PRB}$ is the corresponding risk at policy relevant background for $O_3$. As discussed at the Panel meeting, a preferable alternative is to compute $\Delta = R_{0.8} - R_x$ for various concentrations x. This form would allow focus on the question "What is the difference in the expected number of health effects that will occur at various concentrations of $O_3$, relative to the current standard of 0.08?" A key advantage of $\Delta$ is that it does not depend on the choice of PRB, and thus is free of the uncertainties surrounding estimation of PRB.

## Dr. Henry Gong

Post-Meeting Comments on Chapter 6 (Staff Conclusions on Primary O3 NAAQS) in Second Draft of Ozone Staff Paper (July 2006).
Henry Gong, Jr., M.D.
August 30, 2006

I did not present pre-meeting comments for this chapter. The discussion of Chapter 6 by the Committee as a whole helped me with my evaluation of this chapter. My comments about this important chapter are as follows:

1. I believe that most of my comments regarding Chapter 5 also apply to Chapter 6, especially in terms of the foundation for Staff decision-making about clinical endpoints and their public health relevance. The essential points are that many clinical endpoints (e.g., see Figure 3-4, page 3-53) may not be sufficiently evaluated with quantitative risk assessment but nonetheless constitute highly relevant evidence-based ozone-related health effects. Such evidence and its consistency "trumps" models, according to many clinicians.
2. The "adversity" of the health effects may be described using guidelines developed by the American Thoracic Society for healthy people and people with respiratory impairment (see Tables 3-2 and 3-3, pages 3-70 and 3-71). These tables may be somewhat misleading when comparing exposure responses of healthy people versus respiratory patients. People with pre-existing respiratory disease have varying pulmonary reserves; small percent decrements in lung function may result in more symptoms, interference or limitation of function, use of medications, etc., than anticipated. Even pulmonary patients with "mild" or "moderate" responses may require significantly more therapy (e.g., medications) and/or emergency department visits or hospitalization. The exposure to ambient ozone is typically recurrent (since ozone episodes generally persist over several days), and adverse effects are likely experienced repeatedly in sensitive groups.
3. The weight of the scientific evidence that was compiled and summarized in the Ozone Criterion Document and Ozone Staff Paper clearly indicates that the current ozone standard is inadequate to protect human health. Although more data is desirable such as at low ozone concentrations, we see adverse health effects below the current 0.08 ppm-standard. My reading of the numerous tables and figures indicate that thousands of sensitive children will continue to experience ozone exposures of concern and resulting lung function decrements (and other health effects) at or below 0.07 ppm. [I note that the use of percentages can be confusing and misleading in many of the tables (see Dr. Lianne Sheppard's Final Comments, 8/28/06).] Thus, my overall gestalt is to support a reduction of the 8-hour standard to 0.060-0.070 ppm (to three decimal places) to achieve greater protection of public health. I am still concerned that even this range does not build in an adequate margin of safety for sensitive groups, such as asthmatics. Thus, I favor a standard in the lower end of the proposed range. I anticipate that future research studies will demonstrate evidence of adverse effects in sensitive groups at lower ozone concentrations.

Post-Meeting Comments on Chapter 5 (Characterization of Health Risks) in Second Draft of Ozone Staff Paper (July 2006).
Henry Gong, Jr., M.D.
August 28. 2006

My pre-meeting comments (dated 8/22/06) remain active. The discussion by the Committee as a whole helped me extend my evaluation of this chapter. My further comments are as follows:

4. As I discussed during the meeting (August 24-25), I am concerned that a great deal of emphasis is given to the quantitative risk assessment of lung function decrements in school-age children. The concern about this vulnerable group is appropriate and well supported. Nonetheless, I conclude from this chapter that the emphasis is disproportionate in the sense that numerous other clinical endpoints are down-graded or considered less than optimal, and, thus, should not be effectively or appropriately used to support the adverse health risks from contemporary ozone exposures. These ozone-related endpoints are respiratory symptoms among asthmatic children, increased medication usage, emergency department visits and hospitalizations for respiratory illnesses, and non-accidental, cardiorespiratory deaths. The discussion is up and down regarding their respective causality by ozone, sufficiency of exposure-response data, data consistency, etc. I cannot see how such a large volume of supporting evidence of clinically relevant information can be masked and not even considered in some type of qualitative risk assessment. Figure 3-4: "Effect estimates (with 95%) for associations between short-term ozone exposure and respiratory health outcomes" (Chapter 3, page 3-53) is the best overall visually convincing clinical summary of this argument. This figure should be prominently and effectively used in subsequent communications regarding the adverse respiratory effects of ambient ozone exposure. I believe that the clinicians (physicians) on the CASAC Ozone Review Panel recognize and agree to this point as well. As someone stated during the meeting, the formal risk assessment process is only a piece of the evidence and not the entire argument for ozone-induced adverse health effects.
5. Lung function at low ozone concentrations (0.04, 0.06, and 0.08 ppm): This type of clinical data is very important to obtain and understand. I appreciated and supported the inclusion of Dr. Adams' data into the Ozone Criterion Document and Staff Ozone Paper. Nonetheless, his data (n=30 subjects) must be carefully understood and interpreted. The probabilistic exposure-response relationships for FEV1 decrements (Figures 5-2a,b,c) show a logistic modeled sigmoid curve that is weighted by relatively few data points at the lower ozone concentrations. Are the FEV1 responses normal variations or real ozone responses? Are these responses reproducible in the same subjects? How confident can we be at this time about the lower risk estimates for lung function response with the limited data points? Thus, I am expressing cautious review and interpretation of the small amount of lung function data at these low ozone levels.
6. Page 5-56. The "ozone season" in different study areas appears to be a "moving target" and is not identical from region to region. I am concerned here that we are not comparing comparable ozone seasons. Changing meteorology is another influential factor that alters ozone levels from year to year as well as spatially.
7. A recently published paper (Triche EW et al: Low-level ozone exposure and respiratory symptoms in infants. Environ Health Perspect 2006.114:911-916) did not make it into the

current CASAC ozone review process. Nonetheless, I would like to call attention to it as another indication that respiratory symptoms occur at or below the current EPA Standards.

---

Pre-Meeting Review Comments on Chapter 5 (Characterization of Health Risks) of Second Draft of Ozone Staff Paper (July 2006)*
Henry Gong, Jr., M.D.
August 22, 2006

Responses to Four Queries from Dr. Martin et al. (OAQPS):

1. I generally found the assessment, interpretation, and presentation of the results of the revised exposure analysis in the two documents to be technically sound, balanced, and adequately communicated. The progression of discussion appears logical. I specifically found the presentation on exposure-response functions (section 5.3.1.3) to be reasonable.
2. The set of health endpoints and concentration-response and exposure-response functions used in the risk assessment appear appropriate, given the limited information available for other potentially important endpoints, e.g., airways hyperreactivity, respiratory hospital admissions. As such, it seems that we are underplaying and understating the biological and health impacts of the non-analyzed endpoints of ambient ozone exposures.
3. As far as I can tell, the methods used to conduct the health risk assessment are technically sound and appropriate. I look forward to hearing other reviewers comment on the methods used.
4. The major uncertainties associated with the health risk assessment appear clearly and appropriately characterized.

Other Comments about the Documents:

1. The value for policy relevant background such be stated early to set a reference point.
2. Page 5-12/line 3: "FEV1i." must be a typo?
3. Why are there negative numbers for incidences of mortality (3 columns) in Table 5-11, page 5-51?
4. I cannot find Figure 5-12 (see page 5-71).
5. I do not understand the significance of the statement that "the 8-hr maximum concentrations are on average twice the 24-hour average level." I read it on page 5-73 (lines 37-38) and I believe that it was also stated previously in the chapter.

* Also includes supporting Chapter 3 (Assessment of Risk Based on Controlled Human Exposure Studies) of the Draft Ozone Health Risk Assessment for Selected Urban Areas: Draft Report (June 2006).

# Dr. Paul J. Hanson

Dr. Paul J. Hanson
August 27, 2006

**Review of Chapters 7 and 8 of the
Second Draft OAQPS Staff Paper on Ozone: Policy Assessment of Scientific and Technical Information
and the
Technical Report on Ozone Exposure, Risk, and Impact Assessments for Vegetation: Draft Report July 13, 2006**

**Final Comments from Dr. Paul J. Hanson: August 27, 2006**

The staff has done a credible job of summarizing the important issues of crop and forest tree species responses to tropospheric ozone. Statements of the uncertainties surrounding their conclusions are presented, but some improvements could be made. I do, however, recognize that based on the nature of the available data the statements of uncertainty will default to qualitative statements in some cases (e.g., projected mature tree responses).

## A. Responses to the Panel's Charge Questions on Chapter 7

1. To what extent is the presentation of evidence drawn from the vegetation effects studies assessed in the $O_3$ AQCD technically sound, appropriately balanced, and clearly communicated?

> I was generally pleased with the technical content, balance, and presentation of scientific details contained in the second draft staff paper and the associated technical documentation. Comments and suggested changes for specific items are detailed after the responses to the charge questions.

2. What are the views of the Panel on the appropriateness of staff's weight-of-evidence approach which assesses information from across the various vegetation-related research areas described in the $O_3$ AQCD, including chamber and free air exposure crop yield and tree seedling biomass experimental studies, foliar injury data from biomonitoring plots, and modeled mature tree growth?

> I support this approach and was impressed that a variety of disparate methods could be used to develop a consistent picture of the sensitivity of vegetation to ozone uptake. I have listed a number of concerns that the Staff should consider in the preparation of a final document.

3. To what extent are the methods used to conduct the exposure assessment and the interpretation and presentation of the results of the exposure assessment in the second draft Chapter 7 and the draft Environmental Assessment technical support document technically sound, appropriately balanced, and clearly communicated?

The methods are appropriate to the appointed tasks, but my comments point to areas where details were missing and where the clarity might be improved.

4. To what extent are the uncertainties associated with the exposure analysis clearly and appropriately characterized in the second draft Chapter 7 and the draft Environmental Assessment technical support document?

Uncertainties related to the calculation and enumeration of exposure values and statistics are covered in the document (Page 7-29), but more attention should be paid to the uncertainties of the biological input data with respect to how well they characterize the full breadth of genetic variation within crops and natural species. The document does a reasonable job of characterizing the variability of the available data for cultivated crops (pages 7-42 and 7-43), however, no similar discussion is provided for the response of tree species and the vegetation of ecosystems.

The document should acknowledge that available studies on non-crop species target a combination of species of commercial importance and species known to be sensitive to ozone. That is, much of what we know about the response of non-crop-plant response to ozone results from work to understand why particular species or within-species genotypes exhibited their obvious sensitivity to ozone. The fact that they were sensitive to ozone caused them to be studied more than species that did not exhibit measurable responses. The document should indicate that we have an inadequate quantitative understanding of the full range of ozone sensitivity in natural vegetation. Available data could represent either an over or an underestimate of vegetation response throughout the United States. A dearth of information exists for key species present in many areas. Research is needed to understand the true range of ozone sensitivity of species and within-species genotypes distributed across the landscape.

5. To what extent are the uncertainties associated with the vegetation risk assessment clearly and appropriately characterized in both the second draft Chapter 7 and the draft Environmental Assessment technical support document?

No comment.

6. Staff recognizes that gradients can exist between $O_3$ levels measured at monitor probe heights and those measured over low vegetation canopies. What are the Panel's views on the appropriateness of applying a single adjustment factor to hourly monitoring data to account for the range of potential gradients that can exist across sites and crop and tree seedling canopy structures? Are there alternative approaches or adjustment values the Panel would suggest? Are staff's planned sensitivity analyses appropriate and sufficient?

This issue is discussed in more detail and with much greater clarity within Chapter 7 (pages 7-47 to 7-48) than in the supporting technical report. The authors did a good job of characterizing the desirability of adjusting measured ozone concentrations to relevant crop heights and they clearly recognize that such a procedure "Ideally should account for

the exact height of each monitor, canopy roughness for each crop, and the seasonal and diurnal nature of turbulence." Nevertheless, in their attempt to adjust ozone exposure levels for monitor vs. crop heights they may have improved the reality of ozone exposure for some crop species but degraded the applicability for others (e.g., tall maize crops). Correction factors for crop species growing in windy environments with lower boundary layers would be different than correction factors for crops growing in areas with more stagnant air masses.

I do not suggest that Staff eliminate the completed analysis, but rather include an additional analysis based on the POES data without the 10% adjustment. Having both data sets available will provide a bracket of responses within which the reality probably lies for all crops. Based on their presentation at the public meeting, I see that the Staff has indeed begun to make these calculations, and understand that they will be included in the final document.

7. To what extent do the figures aid in clarifying the text? Should more or less information of this type be included in the final Chapter 7 or its Appendices?

The figures were generally appropriate and helpful, but I made some recommendations on specific figures and tables in the detailed comments listed below.

8. Given the lack of quantitative information on $O_3$-related ecosystem effects, what are the Panel's views on the appropriateness of how this topic is addressed in the second draft Chapter 7?

The document provides a concise summary of the important issues to be considered at the ecosystem scale, but appropriately concludes, "it is difficult to quantify the contribution of $O_3$ due to the combination of stresses present in ecosystems".

**B. Responses to the Panel's Charge Questions on Chapter 8**

1. Does the Panel generally agree that the secondary standard options identified by staff (including indicator, averaging time, form, and level) are generally consistent with the available scientific and technical information and are appropriate for consideration by the Administrator?

Given the reality of the available data and the status of our understanding of the response of vegetation to ozone exposure and uptake, the proposed indicators, averaging times, and forms are appropriate. I recognize that specifying the level or levels to be applied to the new form of the secondary standard is a difficult task underscored by substantial uncertainties in the data. Nevertheless, I encourage the Staff to propose a level specific to the protection of vegetation independent of the primary standard. This is an important distinction. By providing an independent justification for the averaging time, form and level for the secondary standard the Administrator will have the option of considering a change to the secondary standard even if the primary standard were to be left as is.

The Staff's recommendation for changing the form of the secondary standard and a range of proposed levels should be accompanied by documented statements of the benefits that society would gain by achieving the proposed standard. Tangible benefits that might be calculated include increased crop yield, forest production, and tree seedling survival. Conclusions regarding the benefits to be gained by a change in the secondary standard should be accompanied by acknowledgement of the uncertainties involved. The resulting summary would demonstrate to policy makers and the public the utility of considering a new secondary NAAQS for ozone.

In the future, alternate approaches based on mechanistic models of plant and ecosystem ozone uptake and subsequent response can and should be developed and discussed. However, the application of such detailed approaches to the combination of species and biomes present across the United States is not possible with current data (see page 7-4).

**C. Specific comments and editorial suggestions by document**

In the following section I provide comments on the items that I questioned or for which I would like to point out possible editorial modifications. Readers of these comments should not view the list as a negative reflection on the document as a whole. Overall, I found that Chapters 7 and 8 and the associated technical document contained many excellent conclusions and statements. I have simply not taken the time to enumerate and point out the positive.

**Chapter 7**

Page 7-7 lines 21 to 30: The authors fall into the trap of using teleological concepts to describe active rather than reactive plant responses to stress. Plants do not think about responding to ozone stress (they can't). On line 22 the phrase "plants have a variety of compensatory mechanisms" should probably be changed to 'plants exhibit a variety of compensatory mechanisms'. What looks like or appears to be variable repair capacity among species or genotypes could be an induced response to wounding, but it could also be the result of inherently different membrane or protein turnover within those same tissues.

Page 7-7 line 4: One might replace "reallocation" with 'translocation'.

Page 7-8 last paragraph: Please reword as "…with the exception of a few clones of the genera *Populus*, which can be highly sensitive and in some cases are as sensitive to O3 as annual plants." Some *Populus* clones are also very insensitive to ozone.

Page 7-9 line 9: I would prefer that "cause a" be changed to 'result in a'.

Page 7-9 line 32: Replace this line with '…in vigor may be plant death for sensitive plant species'. The response discussed would not be general to all plants.

Page 7-12 line 16: Change "…which diseases are likely…" to '…the disease-ozone interactions likely…'.

Page 7-12 line 27: Would it be appropriate to indicate that the competitive response is driven indirectly by changes in plant species ability to compete for water and nutrients?

Page 7-13: Change "increasing ozone flux" to 'foliar ozone uptake'.

Page 7-14 line 16: The authors might add a new heading at this location: **Climatic Change**.

Page 7-14 lines 24 to 28: The authors could consider citing a few modeling papers that have dealt with offsetting $CO_2$ and ozone responses.

Page 7-14 lines 30 to 34: Suggested rewording: '…simulations will be important in specifying hypotheses to be tested for the many complex interactions of ozone and various combinations of environmental factors. The results obtained will, of course, only be as reliable as the input data for their parameterization, and additional data from organized, systematic study will be needed to judge the efficacy of such model runs.'

Page 7-15 lines 1 to 3: A reference or the name of the EPA program should be provided.

Page 7-16 line 25: Delete the word 'plots'.

Page 7-19: Figures 7-1 and 7-2 have inadequate captions. They don't stand on their own. I was unable to follow these figures until they were explained at the meeting.

The caption for Figure 7-3 was inadvertently split by the figure and "Smoke" should be 'Smoky'.

Page 7-28 line 26: What are the criteria for the statement that the current NAAQS may not provide adequate protection at this point in the text? At this point, all that has been discussed are the exposure metrics. They have yet to be connected to biological response or a discussion of risk. The authors are associating attainment of ozone atmospheres above the current standard with a negative biological response. Although this may be true for some plant species (perhaps even many plant species), it won't be true for all plant species. I would reserve the statements regarding levels of protection for later discussion.

Page 7-41: Exposure response information can be developed from FACE exposure systems, but one must recognize the potential for significant gradients of exposure gas concentrations throughout the exposure rings. While the FACE protocols minimize exposure concentration gradients, plants near the gas emitters will be exposed to larger concentrations than centrally located plants near the air monitoring/sampling point. In the case of $CO_2$ studies this has often been considered a minor point, but in the context of ozone response studies where peak concentrations are known to disproportionately affect plant function, subtle gradients could be very important. If growth data are averaged across the entire ring, the mean response associated with ozone concentrations measured only at the center of the ring will lead to overestimates of plant response when extrapolated directly to other field locations.

Page 7-43: The authors should delete the comparison of the SoyFACE results with the CR-functions. The hail damage totally confounds the desired comparison to the point that the

comparison should not have been attempted. The Morgan et al. (2006) article is missing from the reference list.

Page 7-44 figure 7-13 versus Figure 7D-1 in the annex: When response surface data such as these are interconverted, how is it done? That is, how exactly is the x-axis based on 12 hour SUM06 altered to derive the exact or approximate equivalent for the "4$^{th}$ highest daily 8-h average"? Is an approximate multiplier applied or are the values recalculated from raw hourly ozone values collected at the time of the original experiment? I asked a similar question at the public meeting and wasn't sure that I got my message across. I hope that this alternate wording of the question helps.

Page 7-51 line 17: Should "Taylor et al. 1993" be listed as 'Taylor 1993'.

Page 7-57: What *Populus* clonal data were used as the base response function for the development of this map? Can the authors characterize the extent to which that response function is appropriate and applicable for the full range of *Populus* distributed nationwide?

Page 7-58 to 7-60: This section starts out with so many qualifying statements about the limited use of foliar injury responses I began to wonder why the authors bothered to include it in the document. Further reading of this section, however, suggested that the FIA foliar injury data has real value as qualitative indicator of ozone exposure. If carefully measured and interpreted in the context of available ozone concentration measurements, I believe foliar injury data may have found a useful role as a semi-quantitative assessment of exposure. The authors do an excellent job of characterizing the limitations of this analysis. Such data will likely never be refined enough to help characterize appropriate levels for future standards, but they demonstrate nicely the relative magnitude and regional nature of vegetation responses to ozone.

Appendix 7A Page 3 and 4:

I believe that item number one leaves the reader with an inappropriate conclusion. The real issue is that for flux-based indices to work they would need to be based on perhaps hourly time step models to ensure that ozone concentration and foliar conductance data are appropriately matched. I don't believe that a flux-based approach can be reduced to a simple mathematical index. The full mechanistic model based on a 24-hour calculation would be essential for success. In addition, a flux based model of ozone response would also need to include model components necessary to account for the transient development of soil water deficits that would lead to stomatal closure and result in reduced ozone uptake by foliage.

Appendix 7D Page 16:

Tables 7E-1, 7E-2, 7E-3 and similar Table in the technical report (pages 5-6 and 5-7) would be more useful for future reference if the parameter estimates for each regression analysis (mean and perhaps max and min) were accompanied by their respective standard deviations. Future probabilistic estimates of response, as apposed to the deterministic approach provided in the tables, would be made possible by the inclusion of these estimates of parameter error. In fact, if only the parameter estimates for those curves characterizing the mean response data were provided with the appropriate estimates of error, calculated max and min estimates could be reproduced for defined confidence limits around the mean.

Providing such data in these key documents could very well be the basis for new and independent analyses by other researchers.

**Chapter 8**

Page 8-5 lines 4 to 16: The predicted value of anticipated loss would be greater if the ozone concentration reduction of 10% had not been applied to the POES data prior to this analysis. For reasons stated earlier, one might consider these estimates of loss to therefore be too low. Please include a discussion of the unadjusted values as presented by Staff during the public meeting.

Page 8-11: Please include an explanation and justification for why the cumulative exposure indices are limited to a 3-month window? Vegetation grows for much longer than 3 months in many regions of the United States. There may be an obvious answer apparent to those associated with the standard setting process, but I believe it should be spelled out in the Staff paper.

Page 8-14: As discussed and concluded during the public discussion, I agree that the Staff paper should not recommend keeping the secondary standard identical to the primary. Such an option remains a default that the Administrator could consider. It need not be listed and certainly is not supported by the weight-of-evidence.

Research needs:

All of the proposed research needs listed on pages 8-14 to 8-16 are justified, but they are not prioritized and therefore represent a simple wish list of desired research. To move ozone effects research ahead at an optimum pace, any list of proposed research needs focus, prioritization and integration. Limited financial resources underscore the need for making tough choices regarding the order of completion of research objectives. A long wish list is a very scary budget prospect, but a prioritized research program to be addressed over time represents a palatable option for sustained progress in times of limited financial resources.

My personal opinion is that baseline mechanistic research for the identification of the role of antioxidants and tissue replacement (i.e., repair) in conveying species or species genotype specific ozone resilience has received inadequate attention in recent years. The physics of ozone transport to and into foliage is reasonably well understood, but improved models may be required to characterize this phenomenon within complex canopies of mixed vegetation. Development and implementation of a plan to characterize the full range of natural plant sensitivity to ozone for plant species with important commercial and/or aesthetic value to society is needed. Such data will ultimately be needed to translate plant specific responses to the monetary or intangible values they have throughout the United States.

Notwithstanding excellent past suggestions for research to improve health and ecological risk assessments for ozone (EPA December 2001), the EPA has made very limited investment in ozone research since the last CD was produced. As I understand the situation, reduced emphasis on ozone was driven by the need to refocus funds and effort on other criteria pollutants (e.g., PM). To advance the science in time for the next periodic evaluation of the Ozone and Related Photochemical Oxidants Standards, however, a fundamental change in the manner in which EPA funds research must be considered. Because such a plan must be realistic and consider the distribution of limited fiscal resources, I recommend that EPA consider a two phase research concept consisting of sustained funding for mechanistic investigations whereby ozone damages

tissues and their functions, to be supplemented periodically by field-scale, exposure-response campaigns to provide tangible empirical data for testing models of plant response to ozone exposure. The timing of new expensive field campaigns could be determined or initiated when major shifts in cultivar use are evident, are currently needed for many dominant tree species, and might be applied to other ecosystem indicator species as their relationship to ecosystem net production is better understood.

Whether my specific recommendations are followed is not important. It is critical, however, that EPA make plans for sustained investment in criteria pollutant research. Without sustained research investment we stand to lose critical institutional knowledge both within and outside EPA necessary to make the periodic evaluation of NAAQS work. Without new research results and sustained monitoring of indicator pollutants at relevant scales the legislated requirement for periodic evaluation of air quality standards could become a pointless exercise.

**Technical report**

The draft *Technical Report on Ozone Exposure, Risk, and Impact Assessments for Vegetation* dated July 13, 2006 contains useful information in support of Chapters 7 and 8 of the Staff paper.

Page 5-7: The 10% downward adjustment of ozone values is not fully explained in this document, but appropriate text for this issue is included in Chapter 7. If the technical report is intended to act as a stand-alone document as well as a supplement, it should contain some of those details.

I personally found Figures 5-5 to 5-9 to be very instructive regarding the yield losses for crops and one or more of these figures might be moved forward into Chapter 7. The analogous figures for tree seedling responses were also very useful (Figures 5-10 to 5-14). Including such figures in Chapter 7 would provide the policy makers with visual information about the species-to-species variation.

The mature tree growth discussion on pages 7-1 and 7-2 should be expanded to include the details for the western tree species simulations discussed in Chapter 7 (i.e., ponderosa pine).

# Dr. Jack Harkema

## Comments on Second Draft of the Ozone Staff Paper
## Chapter 3, "Policy-Relevant Assessment of Health Effects Evidence"

Dr. Jack R. Harkema
CASAC Ozone Review Panel
August 25, 2006

General Comments:

In general this chapter is very well written and adequately summarizes the health effects evidence presented in Chapters 4 through 7 of the CD. The authors have done a good job of integrating the most relevant health-effects data from epidemiologic and human/animal controlled exposure studies in this summary chapter. Emphasis on new reports in the peer-reviewed, scientific literature regarding subpopulations with elevated susceptibility to ozone exposures, effects of ozone exposure on the cardiovascular system, and epidemiology studies concerning the effects of acute exposure of ozone on mortality are most appropriate and adequately presented in this document. The new data presented in the CD and summarized in this chapter is strongly supportive of the current ozone standard, but also suggests that the current standard may not be adequate to protect susceptible populations such as those with cardiopulmonary disease, children and the elderly from adverse health effects. This needs to be more strongly and clearly stated in this chapter.

The revised summary tables and figures in body of the text and appendix are excellent and nicely complement the text. For the most part, more interpretation of the results of key studies has been provided in this revised draft compared to the previous draft. There are, however, areas in this chapter that still need more interpretation of results to help the reader understand the plausible reasons underlying the principle health-effects findings. For example, additional statements are needed on biological mechanisms that may be responsible for seasonal and individual variations in response to ozone exposure in the subsections of 3.4. Numerous animal and human controlled exposure studies have documented that individual exposure history to ozone (or other oxidant pollutants) determines an individual's response to short-term exposure. A person living in an ozone-polluted city will likely develop some form of tolerance to ozone and may respond differently than a person who lives in a relatively clean city and who has not built up a tolerance to ozone. The latter person may be at greater risk for developing adverse health effects in response to acute ozone exposure than the former person. This concept needs to be presented more clearly in the document and how it relates to the issue of threshold determination.

Specific Comments:

| | |
|---|---|
| P3-1, L29-31 | Authors need to state the key findings not just what was examined. |
| P3-6, L7 | Define "BSA" and "triangular exposure." |

P3-15, L15 Epithelial cell proliferation (hyperplasia) is a more prominent and well-documented ozone-induced CAR lesion after acute exposure than is intramural fibrosis.

P3-21, L13 Monkeys exposed to 0.15 ppm (not 0.2) ozone developed airway remodeling (Harkema et. al., Am J Pathol. 1993 Sep;143(3):857-66).

P3-23-24, L35-2 This is an example of "tolerance" in an animal study.

P3-44-45, L34-2 Study by Vedal et al. is a good example of a population with a low level of oxidant air pollution that have a significant association between exposure and mortality at a low ozone concentration (see general comments).

P3-46, L21 Change "inflex" to "influx."

Overall, Chapter 3 is excellent and provides strong supportive documentation for setting the ozone NAAQS.

## Dr. Philip K. Hopke

Comments on Chapter 2
P.K. Hopke

Ozone is of key importance in the atmosphere because it is the easily measured indicator of photochemical oxidation and oxidative chemistry. This chemistry leads to particle formation and the production of oxidize products that serve as irritants (acrolein, carbonyl compounds) and materials that can produce other effects. Thus, this chapter should take the opportunity to indicate the broader role of ozone and related photochemical oxidants on the overall production of air pollution in both the gas and particle phases.

I still believe the role of other ambient photochemical oxidants in producing both health and ecosystem effects has been underestimated in this document. The short rationale provided in Chapter 2 for using ozone as the indicator basically dismisses the role of other oxidants, but clearly is written with the concept that the other oxidants are in the gas phase. There should be at least a mention of the role of ozone in particle formation and the formation of particle bound ROS. We currently know little about the concentrations and potential health effects of these species, but a placeholder for future better understanding of this issue should be added through a few sentences addressing this issue.

One aspect of ozone chemistry and exposure that has not been noted in the staff paper is the formation of indoor particles from ozone that has infiltrated into the indoor environment reacting with organic compounds of indoor origins. I suggest a paragraph something along the lines of that which follows be added to Chapter 2.

Formation of new PM by nucleation has been observed indoors as a result of infiltrated ambient $O_3$ interaction with terpenes (Sarwar et al., 2003; Weschler, 2004). One class of compounds that are active ingredients of indoor air is terpenes (Weschler and Shields, 1997). Terpenes such as α-pinene, and β-pinene are also found in emissions from building material such as pine, hardwood, southern pine etc., (Baumann *et al.*, 1999). The terpene system (Fan *et al.*, 2003; Weschler & Shields, 2003; Weschler, 2003; Sarwar et al., 2003) has been studied and known to produce aerosols with ozone levels normally observed indoors from exchange with outdoor air. Thus, another route of exposure to air pollutants is indoor exposure to particles resulting from infiltrated ambient ozone.

Baumann, G. D. M., Batterman, S.A., Zhang, G., Terpene emissions from particleboard and medium-density fiberboard products, *Forest Products Journal* 49, 49-56, 1999.

Fan, Z., Lioy, P., Weschler, C.J. , Fielder, N., Kipen, H., Zhang, J., Ozone initiated reactions with mixtures of Volatile organic compounds under simulated indoor conditions", *Environmental Science and Technology* 37, 1811-1821, 2003.

Sarwar G, Corsi R, Allen D, Weschler C, The significance of secondary organic aerosol formation and growth in buildings: experimental and computational evidence, *Atmospheric Environment* 37, 1365-1381, 2003

Weschler CJ, Indoor/outdoor connections exemplified by processes that depend on an organic compound's saturation vapor pressure, *Atmospheric Environment* 37, 5455-5465, 2003

Weschler, C. J., Chemical reactions among indoor pollutants: what we've learned in the new millenium, *Indoor Air* 14 (suppl 7): 184-194, 2004.

Weschler CJ, Shields HC, Potential reactions among indoor pollutants, *Atmospheric Environment* 31, 3487-3495, 1997.

Weschler CJ, Shields HC, Experiments probing the influence of air exchange rates on secondary organic aerosols derived from indoor chemistry, *Atmospheric Environment* 37, 5621-5631, 2003.

Otherwise, this chapter provides an adequate background on the atmospheric background for the standard

## Dr. Michael T. Kleinman

Chapter 4: Characterization of Human Exposure

Comments: M. Kleinman

General comments: This chapter is well written however there are some specific points that could be addressed better to strengthen it.

| Page | Line | Comment |
|---|---|---|
| 4-1 | 18 | The criteria for selecting the 12 locations should be summarized or at least a pointer to pg 4-18 (where this is somewhat discussed) should be inserted. |
| 4-2 | 29 | Briefly discuss why dose estimation is not necessary in this exercise. |
| 4-5 | 29 | Is the 180 ppb an average for a room? Perhaps the room size, air exchange and other key parameters could be discussed since neither of the sources cited are primary peer reviewed literature. |
| 4-5 | 30 | A clear statement of the impact of outdoor O3 on indoor O3 concentration should be inserted here to provide the foundation for statements made on pg 4-6, L 3 |
| 4-6 | 3 | The basis for this needs clarification given the earlier statement that people spend most of their time indoors. |
| 4-6 | 10 | Explain why it is not necessary to go further and estimate dose. |
| 4-11 | 31-32 | Is this double-counting, or is the penetration factor computed sans inside source contributions? |
| 4-12 | 31-32 | This could be a problematical concept the way it is phrased. Isn't this moderated by the probabilistic function? i.e the likelihood of one person getting the entire sum of occurrences is vanishingly small. |
| 4-13 | 25 | Among the improvements should be an extension to actual computation of dose. |
| 4-14 | 17 | The degree to which the modeled 12 sites are representative of US populations should be described. |
| 4-15 | 18 | Were the AERs derived from the 12 modeled sites? If not the uncertainty engendered by not matching the sites should be discussed. |
| 4-18 | 3-4 | This should be rephrased since the meaning is not clear. |
| 4-18 | Table 4-2 | The lack of consistency in periods modeled among the sites is a limitation on the generalization of the model. Some additional justification for the partial data coverage should be included. |
| 4-19 | 6 | Change "particularly" to "potentially"? |

| | | |
|---|---|---|
| | 14 | Change " section 4.3.1" to "section 4.3.4.7" |
| 4-32 | Tables 4-8 and 4-9 | There is an apparent logical inconsistency when children in LA and Houston have fewer exposures above 0.07 ppm-8hr cities with lower average O3 levels and fewer children. |

Technical Support Document:

The modeling efforts are well supported in the document. The method by which the attainment scenarios and the calculations of base cases in Chapter 5 is not obvious. There seems to be a logical inconsistency in finding more person-days of excessive exposures in Boston than in Houston and some explanation needs to be provided.

---

Additional Comments
Chapter 2

The issue of policy relevant background should be more extensively described. However, the use of PRB as a modifier for the risk assessments leads to increased uncertainty in the estimates. The PRB is very important in how states will develop implementation plans to achieve the O3 NAAQS, but the issue is very complex. Perhaps it would best be handled on a local level given the key factors such as source of background O3, precursors, reducing agents that may degrade O3 and many others will vary from locale to locale.

Chapter 3
Pg 3-39 L 1-4 Document states: "As discussed in the CD, Section 3.9, using ambient concentrations to determine exposure generally overestimates true personal $O_3$ exposures by approximately 2- to 4-fold in available studies, resulting in biased descriptions of underlying concentration-response relationships and attenuated risk estimates." Given this more emphasis should be given to the large number of epidemiological studies summarized in Figure 3-4 that show clinically important adverse outcomes associated with O3 exposures where most of the exposures are at ambient concentrations BELOW the current NAAQS. This data, in concert with risk estimates in Chapter 5, proves that the current NAAQS does not provide adequate protection and in fact affords no margin of safety. This should be carried forward in Chapter 6 to support a recommendation for setting a standard below the current 0.08 8hr NAAQS.

Chapter 4
General comments: This chapter is well written however there are some specific points that could be addressed better to strengthen it.

| Page | Line | Comment |
|---|---|---|
| 4-1 | 18 | The criteria for selecting the 12 locations should be summarized or at least a pointer to pg 4-18 (where this is somewhat discussed) should be inserted. |
| 4-2 | 29 | Briefly discuss why dose estimation is not necessary in this exercise. |

| | | |
|---|---|---|
| 4-5 | 29 | Is the 180 ppb an average for a room? Perhaps the room size, air exchange and other key parameters could be discussed since neither of the sources cited are primary peer reviewed literature. |
| 4-5 | 30 | A clear statement of the impact of outdoor O3 on indoor O3 concentration should be inserted here to provide the foundation for statements made on pg 4-6, L 3 |
| 4-6 | 3 | The basis for this needs clarification given the earlier statement that people spend most of their time indoors. |
| 4-6 | 10 | Explain why it is not necessary to go further and estimate dose. |
| 4-11 | 31-32 | Is this double-counting, or is the penetration factor computed sans inside source contributions? |
| 4-12 | 31-32 | This could be a problematical concept the way it is phrased. Isn't this moderated by the probabilistic function? i.e the likelihood of one person getting the entire sum of occurrences is vanishingly small. |
| 4-13 | 25 | Among the improvements should be an extension to actual computation of dose. |
| 4-14 | 17 | The degree to which the modeled 12 sites are representative of US populations should be described. |
| 4-15 | 18 | Were the AER's derived from the 12 modeled sites? If not the uncertainty engendered by not matching the sites should be discussed. |
| 4-18 | 3-4 | This should be rephrased since the meaning is not clear. |
| 4-18 | Table 4-2 | The lack of consistency in periods modeled among the sites is a limitation on the generalization of the model. Some additional justification for the partial data coverage should be included. |
| 4-19 | 6 | Change "particularly" to "potentially"? |
| | 14 | Change " section 4.3.1" to "section 4.3.4.7" |
| 4-32 | Tables 4-8 and 4-9 | There is an apparent logical inconsistency when children in LA and Houston have fewer exposures above 0.07 ppm-8hr cities with lower average O3 levels and fewer children. |

Chapter 5

This Chapter is well written and provides a useful description of the exposure characterization process. The chapter does lean heavily on the human exposure data at 0.04 and 0.06 ppm to seemingly justify the findings that there are measurable effects below the currant NAAQS. There is a tendency in the document to downplay the epidemiological findings of important health effects in cities with average concentrations at or below the NAAQS. The point that could be made is that the limiting the studies to days that were below the NAAQS did not change the outcome. Also a more precise description of the fact that the average exposures (used in the epi. Studies) might be below the NAAQS but that at some location an individual sampler could be above (putting the city out of compliance) although very few people in the population would actually have received an excessive exposure, needs to be addressed. A quantitative assessment

of the sensitivity of average exposure to the "peak" exposure might help to put some concrete values on margin of safety.

Chapter 6
Pg 6-2, L18 Document states: "For the purpose of evaluating the level of the $O_3$ standard in this review, we have placed greater weight on U.S. and Canadian studies, taking into account the extent to which such studies have reported statistically significant associations. This is because findings of U.S. and Canadian studies are more directly applicable for quantitative considerations in this review as studies conducted in other countries may well reflect quite different populations, exposure characteristics, and air pollution mixtures." Was the Canadian activities data used in CHAD?

Pg 6-8, L 7-10 Doc. States: "however, when corrected for the effects of exercise in clean air a small percentage (7%) of healthy adult subjects experienced moderate lung function decrements (≥ 10% $FEV_1$) with exposure to 0.06 and 0.04 ppm $O_3$." Some indication that this finding has public health significance in context with the confidence limits around the concentration response function should be added to the paragraph.

Pg 6-11, L 29 Doc States:[in reference to ED visits] " These studies provide evidence of effects in areas that likely would not have met the current standard but do not address the likelihood of effects occurring in areas that likely would have met the current standard." Should be amended to reflect that in several cases, restricting the analyses to days with concentrations below 0.08 ppm still resulted in significant effects on health.

Pg 6-13 to 6-15 Table 6-1: This table is based on the exposure assessment modeling used previously in Chap 4 which provides "artifactual" low values for key cities such as LA and Houston. The low values are due to the non-uniform O3 distribution with specific samplers having exceptionally high values causing an "inflated" design value. When the rollback is performed, the O3 levels for most of the region is pushed to unrealistically low values leading to an inflated estimate of the reduced numbers of cases at the current standard. This leads up to the statement on Pg 6-16 L22 "When 2004 air quality is adjusted to just meet the current 8-hr standard, there are an estimated 50,000 person days of 8-hr, moderate exertion exposures of concern experienced by about 50,000 school age children, which means that there are very few occasions where there are multiple occurrences of exposures of concern." Inclusion of these atypical cities in this analysis has provided a biased view of the value of possible alternative standards. This analysis should be rerun after excluding the "peaky" cities or with some other roll-back scenario that more realistically predicts effects of controls. Another improvement might be to present the percent of the children that still show adverse effects (as presented in Table 5-5) because this would highlight the fact that even when meeting the current standard substantial numbers (amounting to about 10%) of these sensitive children will still experience adverse effects.

Pp 6-23-25, Tables 6-4 and 5. These tables should be expanded to include the alternative scenarios that will be discussed. This is where one could provide some idea of the magnitude of excess effects that could be removed if one adopted a more stringent standard. Some statement

should also be made that these tables show that there is not an adequate margin of safety in the current NAAQS.

Pg 6-28 L31: Doc. States: "Only a few studies presented results for different $O_3$ averaging periods using the same data set. Two of the recent multi-city mortality studies reported associations for multiple averaging times (Bell et al., 2004; Gryparis et al., 2004). Both reported that the effect estimates for different averaging times were not statistically different, though the effect estimates for associations with 1-hr daily maximum $O_3$ concentrations were somewhat larger than those for longer averaging times, especially 24-hr average $O_3$." The 1 hr standard is important for those anomalous cities with greater non-uniformity in O3 exposure. There should be some mention that states such as California have adopted a 1 hr standard to provide a margin of safety not afforded by the 8 hr standard.

Pg 6-31 L10: Doc states "Collectively, the epidemiologic studies are inconclusive." This is inconsistent with the Figure 3-4.

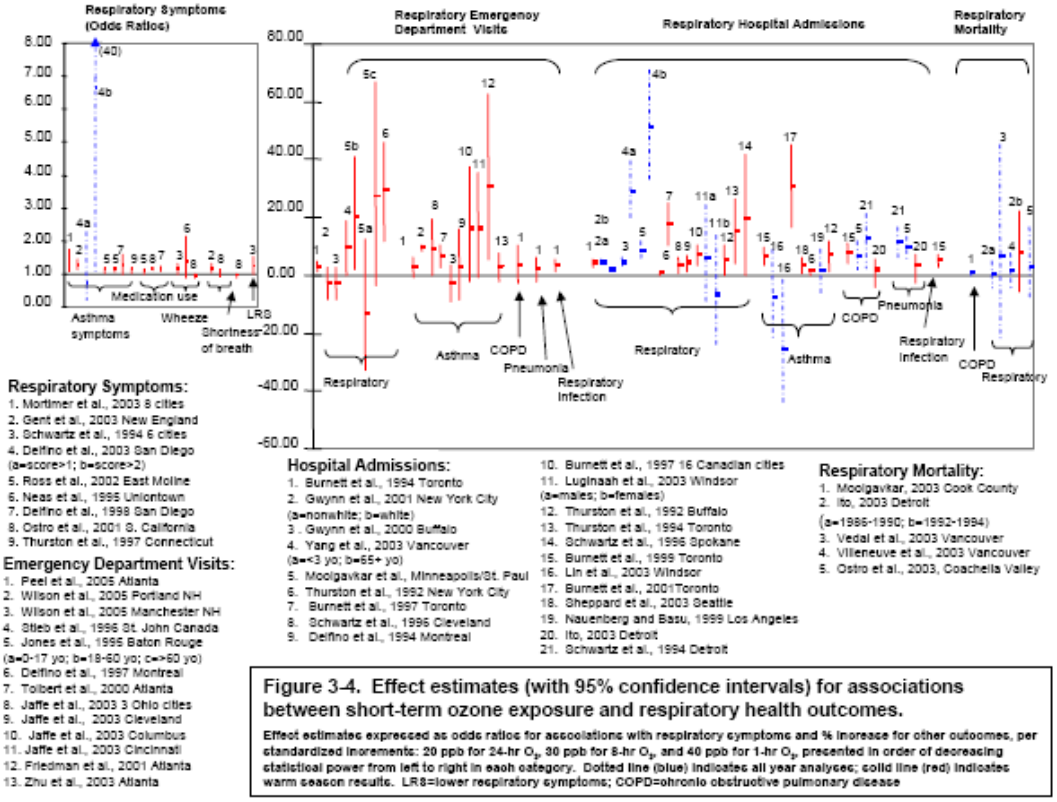


**Figure 3-4. Effect estimates (with 95% confidence intervals) for associations between short-term ozone exposure and respiratory health outcomes.**

**Effect estimates expressed as odds ratios for associations with respiratory symptoms and % increase for other outcomes, per standardized increments: 20 ppb for 24-hr $O_3$, 30 ppb for 8-hr $O_3$, and 40 ppb for 1-hr $O_3$, presented in order of decreasing statistical power from left to right in each category. Dotted line (blue) indicates all year analyses; solid line (red) indicates warm season results. LRS=lower respiratory symptoms; COPD=chronic obstructive pulmonary disease**

3-53

Pg 6-40 L10 Doc states "In conclusion, as discussed in the previous section addressing the adequacy of the current standard, just meeting the current 0.08 ppm, 4th daily max standard substantially reduces estimates of exposures of concern and risks of various health effects" An additional statement that even when meeting the standard substantial numbers of sensitive individuals are still expected to experience adverse effects, hence it is unlikely that the present standard provides the ***required*** margin of safety.

Pg6-44 L 17 Doc States: “Such a view might place more limited weight on evidence of more serious morbidity (e.g., associations with hospital admissions) and mortality effects derived from time-series epidemiological studies” Again referring to fig 3-4 these strongly suggest that there is not an adequate margin of safety in the current standard.

Pg6-44 It would be appropriate to restate the current standard to 3 significant figures which is consistent with the precision of current monitoring devices and which will improve the margin of safety by eliminating “rounding up” to 0.084.

# Dr. Allan Legge

FINAL REVIEW COMMENTS: Allan H. Legge
"Review of the National Ambient Air Quality Standards for Ozone: Policy Assessment of Scientific and Technical Information ( Second Draft ) - OAQPS Staff Paper, July 2006".

Chapter 7: Policy-Relevant Assessment of Welfare Effects Evidence and
Chapter 8: Staff Conclusions on Secondary Ozone NAAQS

Overall Comment:

It is extremely unfortunate that the state of the science regarding ambient ozone and vegetation has not significantly changed since 1996 with the last $O_3$ AQCD. The Agency has essentially done little in the intervening years from 1996 to the present to help solve this problem. The Agency must be strongly criticized for this lack of action. The underlying science to be able to confidently set a meaningful biologically based secondary NAAQS for ozone simply does still not exist. This is especially unfortunate when there is evidence of foliar injury from ozone exposure at the current primary NAAQS for ozone. One can only make scientific advancement with investment and for some reason the Agency has chosen not to make that investment with respect to the science needed to be able to establish a viable and protective secondary NAAQS for ozone.

General Comments:

Despite the above comments, Staff is to be congratulated for attempting such a monumental task. This is especially true given that there has been essentially little new science carried out which has addressed the uncertainties which were evident in the last iteration of the $O_3$ AQCD in 1996. While the presentation of the evidence which Staff has drawn from the current $O_3$ AQCD is as technically sound as the evidence would allow, the large scientific uncertainties remain. Some of these uncertainties are recognized by Staff while others are downplayed as not being important.

Staff, for example, have placed major emphasis on the results of the ozone exposure research experiments carried out in open-top chamber (OTC) under the National Crop Loss Assessment Network (NCLAN) Program for selected agricultural crop species cultivars and for selected tree species seedlings by EPA's National Environmental Effects Research Laboratory-Western Ecology Division (NHEERL-WED) in the 1980's. It has never been shown that the concentration-response (C-R) functions derived from these OTC experiments realistically reflect the response(s) of these plant species cultivars under ambient field conditions.

The argument that one can use the same C-R functions developed during the NCLAN Program on crop species cultivars developed in the middle to late 1970's is highly questionable. This assumes, for example, that the current crop species cultivars in use in 2002 have the same ozone sensitivity 25-30 years later as the crop species cultivars developed in the 1970's. Further, agricultural crop breeders do not directly bred for $O_3$ tolerance, it happens indirectly as a result

of selection for the highest yielding disease resistant plants because the ambient $O_3$ present will be one of the selective factors.

The argument that one can use the same C-R functions developed by NHEERL-WED for selected tree species seedlings grown in pots and exposed to $O_3$ in OTC's and that these data can then extrapolated to mature trees growing under field conditions is highly questionable. Further, one cannot assume that tree species seedlings will have the same sensitivity to $O_3$ as mature trees.

Another very important point is that ozone exposure indices such as SUM06 and W126 which have been used by Staff do not have a biological basis as has been implied. It is ozone uptake by plants which is critical and not the ozone exposure.

While the Staff Paper (SP) is very logical in it's presentation, Staff does not appear to be aware of the very real potential for significant propagation of error and hence significant uncertainty in the analysis prepared for OAQPS by Abt Associates Inc., (2006) [ "Technical Report on Ozone Exposure, Risk, and Impact Assessments for Vegetation: Draft Report - July 13, 2006]. This is a critically important point because the analysis in the Abt Associates Inc. (2006) draft report provides much of the foundation for conclusions drawn by Staff regarding potential secondary standards for $O_3$. Put another way, the analyses provided by Abt Associates Inc. (2006) while quite elegant is essentially a 'house of cards' with an unstable S foundation.

Each chapter builds on the previous chapter with the uncertainties in each chapter being carried forward to the next chapter. These data are then used to generate an economic assessment for agriculture crop loss on a national scale. With respect to the forest trees the seedling data are scaled up to mature trees using the TREGRO Model which includes ozone as a stress factor to estimate yield loss. In the end of all of this one is not sure whether any of this is realistic or that it would be better described as science fiction.

## Dr. Morton Lippmann

M. Lippmann – Review Comments on Second Draft of Ozone Staff Paper, Chapter 2

Specific Comments

| Page | Line(s) | Comment |
|---|---|---|
| 2-47 | 5+ | Policy-relevant background (PRB) needs to be discussed in a much broader context, with recognition that it is influenced by long-range transport and scavenging processes (PRB of 40 ppb over the Pacific Ocean and less over nearby land areas due to reactions with vegetation and with background NO). The uncertainties that PRB creates in apportioning ozone to PRB and North American anthropogenic sources are not relevant to uncertainties in exposure-response relationships at concentrations being considered for a revised ozone NAAQS. |

---

M. Lippmann – Review Comments on Second Draft of Ozone Staff Paper, Chapter 3

Specific Comments

| Page | Line(s) | Comment |
|---|---|---|
| 3-8 | 15-19 | The concentrations that produced the effects described need to be stated. |
| 3-8 | 21 | insert “routine” after “with”. |
| 3-8 | 29-35 | The concentrations that produced the effects described need to be stated. |
| 3-9 | 2 | add ”and that they occurred at concentrations below those used in chamber studies using exercise.” at end of line. |
| 3-10 | 29 | insert “and” before “particularly.” |
| 3-11 | 34 | insert “at rest” after “exposure”. |
| 3-12 | 14 | insert “engaged in moderately high exercise” after “adults”. |
| 3-15 | 24 | insert the duration of the exposure |
| 3-16 | 16 | insert “pattern” after “temporal”. |
| 3-21 | 18 | insert “an association between O3 and” before “acute”. |
| 3-26 | 17 | provide a reference for the cited study. |
| 3-57 | 27 | insert “Short-Term” before “Mortality” |
| 3-59 | 19,20 | change “which” to “that”. |
| 3-61 | 27 | insert “relatively low concentrations of” after “to”. |
| 3-63 | 7 | Continue with a brief discussion of the O3 being part of a mixture. |
| 3-66 | 20 | change “effects” to “mortality”. |
| 3-67 | 2 | change “health” to “lung function”. |

---

M. Lippmann – Review Comments on Second Draft of Ozone Staff Paper, Chapter 4

Specific Comments

| Page | Line(s) | Comment |
|---|---|---|
| 4-1 | 30 | Say something here about other important at-risk groups, such as seniors for excess daily mortality. |
| 4-3 | 12 | change "NOx" to "NO". |
| 4-4 | 11 | add ", respectively" at end of line. |
| 4-5 | 24 | change "NOx" to "NO". |
| 4-13 | 11 | add "elderly" as another population at risk (for premature mortality). |
| 4-16 | 6 | change "until we have sufficient information characterizing" to "insofar as other variables have greater effects, for example" |
| 4-18 | 19 | clarify what the New York urban area covers, specifically how far it extends into PA. |
| 4-30 | 5 | change "city' to "urban area". |
| 4-47 | 9,10 | Sacramento is not an inland portion of the Los Angeles area. |
| 4-47 | 2-25 | The text says Apex was under-predicting. This is a rare case where I would suspect the measurements rather than the model. The measurements were made with ozone passive monitors, which are notoriously sensitive to ambient air velocity across the inlet, and these velocities were likely to be very different across the inlets of the personal, indoor, and outdoor units. |

---

M. Lippmann – Review Comments on Second Draft of Ozone Staff Paper, Chapter 5

Specific Comments

| Page | Line(s) | Comment |
|---|---|---|
| 5-1 | 26 | change "or" to ", or are". |
| 5-3 | 27 | add "It is also important to consider that O3 in ambient air is present in a complex mixture of air pollutants, and that some other components of the mixture may play some role in the health-related effects". At end of line |
| 5-6 | 11 | add "as well as similar responses in outdoor workers and others engaged in recreational outdoor activities" at end of line. |

5-6 14 insert “young” before “adults”.
5-21 7 insert”, but larger than” after “to”.
5-26 28 change “which are not statistically significant” to “that did not yield statistically significant results”.

5-27 22 change “nitrogen dioxide, sulfur dioxide, or carbon monoxide” to “NO2, SO2, or CO”.

---

M. Lippmann – Review Comments on Second Draft of Ozone Staff Paper, Chapter 6

Specific Comments

| Page | Line(s) | Comment |
|---|---|---|
| 6-3 | 4 | insert “of O3 exposures” after “associations”. |
| 6-3 | 15 | insert “O3” before “air” |
| 6-3 | 16 | change “quality” to “concentration”. |
| 6-3 | 22 | change “an” to “a photochemical oxidant”. |
| 6-3 | 32 | insert “and other photochemical oxidants” after “O3”. |
| 6-4 | 12 | insert “indices of pulmonary function, such as” before “forced.” |
| 6-4 | 12,13 | delete “also known as lung function decrements” |
| 6-4 | 24 | insert”, who were judged to be a sensitive subgroup of concern” after “children”. |
| 6-8 | 33 | change “may” to “generally”. |

# Dr. Frederick J. Miller

**Dr. Fred J. Miller**
**August 29, 2006**

**Policy-Relevant Assessment of Health Effects Evidence**
**Chapter 3, OAQPS Staff Paper – Second Draft**

## General Comments

Overall, Chapter 3 now represents a good assessment of the health effects evidence to serve as the basis for characterizing the health risks in humans from exposure to ozone. The material is well written and organized. For the most part, the interpretation of inferences the data may support are reasonable; however, there are a few instances where this reviewer considers staff to have "stretched" to infer the data support their statement (see specific comments below).

Probably my major points of disagreement with the material in this chapter are:

- Discussion of a policy relevant assessment of the evidence for health effects concerning the ranges for changes in FEV1 that are considered to be small, moderate, or large for persons with impaired respiratory systems. Discussion of this issue during the August 24-25$^{th}$ meeting yielded an agreement to not change the table values but to emphasize that a 10% reduction in asthmatics or persons with COPD could have more serious consequences than a similar reduction in a person without a compromised lung.
- The chapter comes across with the view that central monitors in time series epi studies are giving the true picture. Yet we know from material contained in the CD that personal monitors consistently show lower levels. Staff seem to write this off but it has tremendous implications for Chapter 6 relative to the case for lowering the $O_3$ standard and for margin of safety considerations.
- While the discussion of the low level exposures used in the controlled human studies by Adams and colleagues is technically correct that no statistically significant changes were found in FEV1 compared to filtered air, the fact that a reasonable percent of the subjects had large decrements is glossed over. These were healthy subjects, so the percentage of asthmatic subjects, if they had been studied, would be considerably greater. The fact that asthmatics respond with greater differences is shown in other studies discussed in the chapter.
- The lack of statistical power is offered for why there appears to be an inconsistent effect seen for COPD mortality. While the same arguments applied for PM, I don't recall an emphasis on a lack of power. You can't have it both ways – yes for PM and no for $O_3$ using many of the same studies.

## Specific Comments

| | |
|---|---|
| p. 3-15, l. 14 | It is incorrect to claim that short-term ozone exposures cause fibrosis. The wording needs to be changed here. I believe the authors are intending to convey that fibrotic changes occur with short-term exposure. |
| P. 3-17, l. 14 | Consistently, the authors seem to push the false negative possibility for |

| | |
|---|---|
| | explaining discrepancies when the finding they are defending could just as well be a false positive. Better balance here and throughout the chapter is needed. |
| p. 3-18, l.24 | In this section, an Austrian study is cited as showing growth-related increases in lung function over the summer as being reduced with exposures ranging from 32.5 to 37.3 ppb. This study begs the question how low can you go and what is really being measured as growth related increases. This reviewer questions whether confounders are present and whether this kind of study can really be interpreted as growth-related. |
| p. 3-32, l. 6 | The authors state that a positive association was seen between long-term $O_3$concentrations in the warm months, but this association is not statistically significant. The writing comes across biased. If the association had been slightly negative with no statistical significance, I doubt the authors would have written about this. Caution against over stressing weak associations as there is more than enough positive data on other endpoints to carry the day that the current standard is inadequate to protect public health with a margin of safety. |
| p. 35, l. 9 | The rationale and accuracy of this sentence escapes me. Please clarify why the assumption that the chemical reactivity of $O_3$ does not induce strong temporal correlations is needed or is correct. |
| p. 3-47, l. 2 | The Sexton et al. (2004) reference cited here does not appear in the list of references. |
| p. 3-55, ¶ 3 | The authors seem to be stretching how far the data can be interpreted to imply that $O_3$ can have a role in producing effects on the cardiovascular system. The current paragraph is highly speculative and should be rewritten. |
| p. 3-57 | This section on mortality-related health endpoints includes a number of sentences that are a stretch of the available data or that need to be made more factual if they are to be used. For example, on p. 3-58, l. 9 there is a sentence that says the NHNE follow up data have 20% of adults with a reduced $FEV_1$ value that suggests impaired lung function. What is the % reduction being referred to? The reactions of O3 with cholesterol in lung surfactant are stated to generate products responsible for atherosclerotic plaque formation in arterial walls. Where are the data to support this statement? |
| p. 3-61, l. 36 | In the description of the $FEV_1$ change in berry pickers in Canada, the adjective “large” is used to describe a 5% decrease. I seriously doubt that pulmonary clinicians would classify this decrease as large. |
| Table 3-3 | This reviewer does not agree that the same levels of $FEV_1$ decrements as for healthy individuals should be used to describe the degree of change in this variable for those with impaired respiratory symptoms. The statement is made in the text that there is no need to change the criteria from the previous NAAQS review. However, staff acknowledge on p. 3-72 that a 5 to 15% change in FEV1 is considered |

| | |
|---|---|
| | to have clinical importance to asthma morbidity. |
| p. 3-81, l. 14 | Strike "our" from this sentence. |
| p. 3-84, l. 1 | What is meant by this sentence? Provide references for the statement. |

---

**Chapter 4 OAQPS Staff Paper – Second Draft**

There are numerous places in Chapter 4 where only part of the data is presented with the reader being required to go to technical support documents for the "rest of the story". This should not be the case. If the data for one alternative scenario are presented, the data for other scenarios that were considered should also be presented. The number of cases expected for a given health endpoint and a given exposure scenario is helpful but staff need to also describe what percent of the population is represented by these cases. The reader should not have to dig this out of supporting documents.

As part of the sensitivity analyses associated with using the APEX model, staff have provided some information on the number of children experiencing at least 3 or more exposures above various ozone levels for Boston and Huston with changes in the activity database, ozone decay rate, air exchange rate, etc (pages 4-44 to 4-46). This reviewer appreciates finally receiving this type of information, but staff have not gone far enough. One has to go to one of the supporting documents (Ozone Population Exposure Analysis for Selected Urban Areas) to find that, when considering exceedances above 70 ppb, for Boston the number of children having these events with moderate exercise represents 7% of the children and about 1-2% of the children in Houston. Moreover, the figures on pages 50-52 of that document show large numbers of children with 5 or more exceedances.

The frequency of exceedance is at the heart of the question about the level of the standard in order to protect against adverse health effects. This reviewer would submit that a child experiencing only one exceedance to an ozone level slightly above the current standard is not likely to be of significant concern from a public health perspective. However, the same child having 3 to 5 exceedances would be of great concern due to the repetitive injury to the lungs and the increased chance for more serious health outcomes to be experienced. Thus, for every scenario for potential alternative standards, this reviewer wants to see the percentage of children that will have 3 or more, 5 or more, etc exceedances.

**Specific Comments**

| | |
|---|---|
| p. 4-2, l. 28 | The statement is made that staff are modeling intake of $O_3$ and not dose, which is correct. However, staff fail to make use of the dosimetry models that are available. This reviewer has heard that after the last NAAQS review staff were told to forget about using dosimetry models because certain "scientists" did not believe they could be used. Significant improvements in ozone dosimetry modeling have been made over the last decade. This reviewer would be glad to debate the merits of the dosimetry models and their uncertainties compared to the uncertainties present in |

| | |
|---|---|
| | various biological endpoints and epidemiology studies.<br><br>More importantly, by not considering the actual dose of ozone to the centriacinar region, we may well be putting all our eggs in the "asthmatic child basket" when outdoor workers, who inhale much larger amounts of ozone over longer time periods each day and over the entire ozone season may well be the most affected subpopulation. As Dr. Charles Plopper pointed out during the O3 CD discussions, the frequency and duration are important determinants of the type of injury and the probability for the development of serious pulmonary disease. Discussions with EPA staff during the August 24-25th CASAC meeting indicate that there are not sufficient data available to be able to model risk to outdoor workers, and staff agreed to emphasize this point in the final staff paper. |
| p. 4-5, l. 26 | Information is given on how older photocopiers and laser printers run under highly unlikely scenarios can generate 150-180 ppb of ozone. Can staff provide any data on how many such older copiers and printers are still in use? This reviewer highly doubts than many such machines are still in use. |
| P. 4-14, l. 15 | Why do staff contend that a source of uncertainty results from modeling more than one ozone season? Having more seasonal data simply provides data on the variability over years, which is a good thing to know. |
| p. 4-16, l. 19 | The age of the diary information is a significant source of uncertainty for the exposure analysis. About 50% of the diary data are more than 15 years old. This reviewer strongly feels that the extent of computer and video game availability and popularity makes the use of these older data meaningless. |
| p. 4-19, l. 16 | Classifying 45% of the children as "active" seems too high. How much of this classification depends upon the outdated diary studies? |
| p. 4-21, l. 10 | More and more communities are going to year-round schools, so not including school commuting may have more of an impact than staff have supposed. |
| p. 4-47, l. 18 | The APEX model is shown to have a negative bias, which is 30 to 50% most of the time. The paragraph discussing this minimizes the potential impact of this bias. Material in the technical support document should be brought forward for an expanded discussion of the bias. This review considers the bias to be a major weakness in the risk assessment if the implication that significantly more exceedances than predicted are actually occurring due to APEX **always under predicting exposure**. |

# Dr. Maria Morandi

POST-MEETING COMMENTS ON THE 2ND DRAFT OF THE OZONE STAFF PAPER
Maria Morandi

Comments on Chapter 6

Charge Question 1: What are the views of the Panel on the approach taken by staff (as discussed in Chapter 6) of using both evidence-based and quantitative exposure- and risk-based considerations in drawing conclusions and identifying options as to a range of standards to protect against health effects associated with exposure to O3, alone and in combination with the ambient mix of photochemical oxidants, for consideration in this review of the primary O3 NAAQS?

Answer: In general, Chapter 6 delineates clearly the rationale of the staff for the recommendations to be proposed to the Administrator, with appropriate referencing to the relevant sections of the CD and/or Chapters 3-5 of the staff document. Using both evidence-based and quantitative exposure and risk considerations is a reasonable approach. The available scientific evidence since the last CD overwhelmingly supports the staff's recommendation that any new primary standard should be at least as protective as the current one. Clearly, there is no evidence that the current standard is too conservative. The evidence also supports the staff's decision that, while there is some evidence for chronic effects on lung function growth, there is insufficient data for establishing a quantitative standard directed at this chronic effect at the present time.

Chapter 6 elaborates in detail on the additional evidence for effects for which the evidence was very limited at the time the current standard was set - specifically cardiovascular and mortality outcomes and effects on asthmatics, especially children. My sense upon reading this discussion, which reflects the CD but not the Review Panel's general sense on the robustness of these studies, is that evidence for cardiovascular/mortality outcomes is still limited and based on approaches that raise methodological and other questions. There is a contrast between these concerns and the weight the mortality studies receive in this Chapter, and elsewhere in the document, although, in the end, they are not included in the formulation of the various standards.

Charge question 2. Does the Panel generally agree that the range of alternative primary O3 standards identified in Chapter 6 is generally consistent with the available scientific information and is appropriate for consideration by the Administrator?

On balance, the scientific evidence for effects on asthmatic children and the risk analysis in the Draft document does not support the recommendation that the current standard is sufficiently protective of this sensitive population, so that a lowering of the 8-hour standard appears justified on this basis. However, it is not clear to me that the evidence for supporting 8-hour standard would be protective of for this group without consideration of a shorter-duration standard is presented in a convincing manner as indicated in conclusion statement (2) on page 6-43 (also lines 14-22 on page 6-30). Clearly, this statement is appropriate for effects observed in healthy, adult volunteers, but the evidence presented in the document does not shows that it would be as

protective of sensitive asthmatic children without a shorter-duration standard in locations where such excursions occur (i.e., cities such as Houston and LA). There is a somewhat "waving of the hands" feel to the discussion in the Chapter that one side describes the evidence for effects on this sensitive subgroup below the 1-hour or 8-hour standard, but appears not to follow through in the recommendations to the Administrator (either for the 0.08/4 or 0.07/7 ppb/8hr standard). The evidence and risk analysis presented in Chapter 6 do not support keeping the current standard as an option.

Charge question 3. What are the views of the Panel on the key uncertainties and O3 research recommendations discussed in Chapter 6?

Answer: The extensive discussion of the 0.08 ppm vs. 80 ppb concentration appears excessive in light of the uncertainties inherent in the exposure and risk evaluation. The significance of the concentration rounding can only be evaluated in light of an uncertainty analysis that includes both exposure and risk uncertainty, and it should be placed in that context.

Other issues include:

1. Adequacy of ozone as a surrogate for effects from other oxidants: it is correct to say that there is no evidence to support other indicator for airborne oxidants at this time. However, actual data to support this statement are limited very limited (i.e., lack of evidence may be the reflection of lack or research). The lack of data available to support the use of ozone as the only and best indicator for photochemical oxidants can be linked directly to recommendation (3) on page 6-46 under research needs.

2) Addition to research needs: Past and current evidence shows that there are especially more susceptible subgroups among otherwise healthy individuals or those with compromised health such as asthmatics. These subgroups are generally called "responders" but the factors associated with such increased susceptibility remain unknown and, consequently, they may still be at risk even with a more restrictive standard. Studies are needed that can identify the underlying causes that make a subgroup among susceptible individuals particularly sensitive to the effects of ozone (or other agents). It would appear that there are different phenotypes within subpopulations with specific chronic diseases, such as asthma, and that each phenotype may be more sensitive to different specific agents, one of them being ozone. Understanding the underlying characteristics that make, for example, a subgroup of asthmatic children especially sensitive to ozone would provide more robust information for setting a protective standard.

Other general comments

In general, the 2nd draft reflects the scientific evidence presented in the AQCD. There are editorial issues that can be corrected (some of these corrections have been made already) that the Staff can address through a thorough editorial review.

Chapter 2 presents the air quality characterization and analyses in a clear manner that reflects the AQCD. However, there is a strenuous effort that drives the analyses for supporting a higher

PRB. As discussed by the Panel during the meeting, the uncertainties in model estimates of PRB for the 12 cities included in the study remain large undetermined. Ozone levels monitored at Trinidad Head are probably indicative of extracontinental background at that site and perhaps along the western seaboard, but not for areas well within the continent to the eastern seaboard. Quite clearly, the PRB has implications for compliance with the standard and, therefore, is an important consideration, but evaluation of commonality in temporal patterns of concentrations at the Trinity Head site and elsewhere in the continent without more fundamental support basic on basic principles is not warranted for a reactive pollutant such as benzene.

Chapter 3 also presents the available evidence in a manner that reflects the AQCD descriptively, but there is fussiness in the presentation as to 1) which are the more robust studies and end points that will form the basis for supporting the standard, and 2) which is the specific sensitive subgroup the standard is meant to protect. While the AQCD presents the plethora of studies and their results without engaging in a critical review, the Draft document needs to be more purposive and clear on its description of the key scientific findings (as compared to the prior AQCD) that will form the basis for the NAAQS recommendation. As indicated in the review of the AQCD and by Panel members in the review of the 1st and 2nd Draft, the time series mortality studies raise a number of methodological questions beyond just ozone, but they are discussed extensively with limited critical evaluation. Given the methodological uncertainties, apparent consistency of results may not necessarily be evidence for the strength of an association but perhaps reflect the similarity of assumptions inherent in the models (and which are typically never stated in an explicit manner). Given the more robust evidence for effects found in the human chamber studies and the more epi evidence indicating effects on asthmatic children, it would appear that the extensive discussion of the mortality effects and the comparatively speculative nature of the proposed mechanisms are not warranted at this time.

Chapter 4 presents a fairly reasonable assessment of exposures under a variety of scenarios. As with any model, evaluation of internal/external consistency and model predictions is critical, but there are very few personal exposure data available for this purpose.

## Mr. Charles Plopper

Comments Regarding Chapter 3-Charles G. Plopper

This chapter is a very satisfactory summation of the health effects evidence evaluated in Chapters 4 through 7 of the CD. In almost all cases, the description of the materials is accurate and provides strong substantiating evidence for the determinations of the subpopulations with elevated susceptibility that may be of greatest public health interest and the understanding of the health-related effects associated with exposures. There are two aspects of this analysis which are germane to understanding health effects evidence that need some expansion. The first of these is the phenomena of tolerance. What this refers to is that the exposure history for an individual will modify their response to subsequent exposure. This is a scientifically well-documented biological aspect of the response to repeated ozone exposures over a period of time. This will have a tremendous impact on such issues as the determination of threshold in both epidemiologic studies and experimental studies using controlled exposures. It is therefore difficult to address the issue of threshold, because the threshold will vary depending on the air quality of the city in which a study population resides. Seasonal variations in 03 levels in the environment will also modify the determination of no effect level in the same study population depending on the time of the season when testing is done. While this issue is emphasized, especially in section 3.4.2.3, the discussion could use clarification in terms of the biological reasons behind the seasonal variation in responses. This may explain some of the lack of strong positive associations between short-term ozone exposure and respiratory mortality, such as discussed in section 3.5 .2. An example of the impact of exposure history on analysis of response is the Vedal, et al. (2003) study which was carried out in a region with very low 03 concentrations and found significant associations with mortality at very low concentrations. Studies conducted in such low pollution environments may be stronger indicators of the level of risk and the threshold for effect, especially in more susceptible populations.

As an editorial note, in some portions of Chapter 3 there is reference to a lack of data from the experimental animal studies on long-term exposure affects. However, in other portions of the chapter, there is discussion of long-term exposure studies, of which there have been an abundance.

The authors should not be hesitant about making strong inferences from the data available. The new data is accurately summarized in the CD and in Chapter 3. It is more strongly supportive of the current standard than was the data when the previous standard was set and indicates that the current standard may be too high to be protective of susceptible populations such as children and the aging. Further, there is a growing body of evidence, well documented by the CD and summarized in Chapter 3, for cardiovascular health effects which must be taken into consideration for evaluating a new, lower standard.

## Dr. Armistead (Ted) Russell

## Expanded Review of the OAQPS Staff Paper – Second Draft

Armistead (Ted) Russell

The comments below deal primarily with the Second Draft of Chapter 4 dealing with Characterization of Human Exposure to Ozone, as well as the related Ozone Population Exposure Analysis for Selected Urban Areas report, the Analysis of Uncertainty in Ozone Population Exposure Modeling memo, and the Uncertainty in Ozone Measurements Memo.

This document includes prior comments, additional comments after the August 24-25 meeting (***provided in italic***), clarifying prior comments (***provided in italic***), and additional remarks (***also in italic***).

Response to Charge Questions:
1. To what extent are the assessment, interpretation, and presentation of the results of the exposure analysis as presented in Chapter 4 (and in the second draft Exposure Analysis technical support document) technically sound, appropriately balanced, and clearly communicated?
**The chapter is more clear this time around, and is reasonably balanced. It draws on the associated report well, bringing forward the main topics and findings. It should, however, be more clear as to how the APEX model and quadratic role-back are specifically used in the impacts analysis.**

2. Are the methods used to conduct the exposure analysis technically sound? Does the Panel have any comments on the methods used?
**For the most part, they have used appropriate methods. The addition of using 2002 is appreciated.**

3. To what extent are the uncertainties associated with the exposure analysis clearly and appropriately characterized in Chapter 4, the Exposure Analysis technical support document, and the uncertainty memorandum?
**I appreciate all of the work they have done and are doing on the uncertainty assessment. Specific comments are found below.**

4. To what extent is the plan for the remaining uncertainty assessment technically sound? Are there other important uncertainties which are not covered? What are the views of the Panel on sensitivity analyses conducted to evaluate the influence of uncertainties in the exposure analysis?
**They are heading in the right directions, though I recommend that they also conduct an uncertainty apportionment as part of their studies. Try not to get too caught up in details and aspects that are so poorly defined that it will slow down the overall assessment. They have done well to state that they will not include some uncertainties.**

**Remarks on the second draft of Chapter 4**: The second draft of Chapter 4 has responded to many of the comments made on the first draft, and is thus more clear than before. I was pleased to see the reanalysis for 2002 in addition to 2004.

Specific comments:

Page 4-47: They show a bias from APEX, but do not provide information on how that impacts their analyses, and also tend to downplay the magnitude of the bias. 50% is big, and the bias, while smaller in other cases, is consistent. *The bias shown continues to concern me, suggesting that the risks/exposures to levels of concern based upon the exposures found by APEX are low. If staff can address this, it would be good.*

4-46: I might refrain from calling the uncertainty analysis as being results just yet.

***Chapter 2**: I am pleased to see the measurement and design value uncertainty added to this chapter. They should add to this chapter the possible bias in the measurements as presented, noting that this may mean that in past analyses (particularly those depending upon routine monitoring data), the levels of ozone experienced may have been less than measured, thus increasing the risk at lower exposure levels. They should also suggest that future instruments be designed to address this issue.*

*In regards to PRB, given the contentious nature of the issue, the use of the PRB should be discussed in Chapter 2, along with other measures. However, their current approach has been peer-reviewed, and is appropriate. They should, before the next analyses, continue to refine and evaluate their approach to setting the PRB, and how it is used in the assessments. For example, might they use the PRB in the roll-back formula? This could be done by analyzing CMAQ simulations where the boundary conditions are set to those conditions representative of PRB conditions.*

**Chapter 6**: This chapter should deal not only with the recommended standard, but also with the precision of the standard. They have done an analysis of the uncertainty in the observations, and find that most of the uncertainty is due to instrument bias. They have also looked at the impact of the precision chosen on results. It is apparent that the exposure is sensitive to the choice of not only the standard, but also the precision chosen. It would seem unlikely that given the large number of monitors used in the various health studies upon which the standard is based, it is unlikely that the bias identified for a single monitor is appropriate, and this should be reflected in the precision of the standard. This chapter should discuss this issue and make an appropriate recommendation.

*I was pleased to hear that they are planning to suggest that the standard be given to three decimal places (e.g., 0,080 or 0.070 ppm) as this is strongly suggested by their analyses. This should not be just recommended, but strongly recommended.*

*As an engineer, I can not speak so directly about the interpretation of the health data, but they very specifically note that (page 6-16) depending upon the year rolled back to meet the current standard, 20,000-600,000 children are estimated to experience exposures of concern. Thus, the*

*current standard is not protective, and would leave a large number of children at risk. It should be noted that the range should not be interpreted as that the number of children experience exposures of concern will vary depending upon year, e.g., from 20,000-600,000, so maintaining the current standard will lead to a very large number of exposures of concern in some years. On page 6-32, the SP notes that lowering the standard to 0.07 would still have, relative to just meeting the standard, 0-10% of the population experiencing exposures of concern, suggesting that there is a strong likelihood of continued exposure to exposures of concern. Such an analysis argues that the maximum standard in the future should be 0.07(0). No data presented suggest that the current standard is adequate.*

**Remarks on the second draft of Ozone Population Exposure Analysis for Selected Urban Areas:** Again, the draft is more clear: thanks. However, they have not "corrected" some of the mathematical presentation (e.g., the del operator) is still used inappropriately (why?).

Page 2: Sec. 1.2: Update year.

Page 19: Eq. 2-15: Need to correct subscripting.

Page 20: You switch units between hours and minutes in the analysis.

Page 27: I still think it should be "… differentiate **between** people…"

Page 28: Are you able to correlate between meteorology, ozone and activity together?

Page 44, Pare 1: Awkward.

Page 64: Needs a bit more explanation.

Chapter 7: They really need to explain what the low bias from APEX really means and how it will be accounted for. Also, why use ug/$m^3$ in this section, ppm/ppb in others?

Specific Comments:

Page 2: Section 1.2: They still have just 2004 as the application year.

**Remarks on the Draft Analysis of Uncertainty in Ozone Population Exposure Analysis for Selected Urban Areas:**

Overall: This is a nice start, but obviously is not a final product, so the comments are on what is currently available. As a general question, are you planning to attribute uncertainty to specific parameters and inputs? If not, why not? I would strongly recommend conducting a first order (or higher) analysis on contributor attribution.

Page 2: Add "Bias" to the list of important concepts in the list.

Page 8: Paragraph 1 is a repeat.

Page 15: Using the ratio approach will give a "bias" and it is not surprising that the central value is 1.065 (what would appear to be a 6.5% bias). Instead, use absolute values, unless you plan to use the ratio to scale your exposures. The last line of this paragraph is opaque, and they need to come up with a better explanation/reasoning. *I would recommend that they not sue the ratio approach, but use the absolute differences in magnitudes, as this will likely show much less bias, and should lead to less skewness later in their analyses. If they wish, they can calculate the relative bias as the (average absolute bias)/(average concentration).*

Page 16 (et al.): Please label all axis of the graph. Also, the last sentence needs to be better justified. The figure caption for Figure 7 (and similar figures) is insufficient.

Page 15: No need to use additional methods for spatial interpolation… they will give different, possibly slightly better, estimates, (they picked two of the easiest approaches, not best) but the difference will likely be quite minor.

Page 18: Table 7 mixes uncertainty units, likely unnecessarily.

Page 19: Fig. 10: Are the standard deviations available? Is the trend for $O_3$ significant?

Page 21: Is the ratio really 0.12-0.16? This strikes me as a very low number for locations outside of street canyons.

Pages 22- 27: The rollback process and the resulting Figures need to be better explained in the text.

Page 39: If you are really going to do what is shown in Diagram 1, you can do some uncertainty apportionment. Not sure why this would be done just to find influential statistics separate of a complete uncertainty analysis.

Page 43 et seq.: I am not sure I grab why they are saying that averaging together a set of simulations is not as accurate as having a single simulation with the same net number of "people". Also, it is not obvious why you need 60,000 given the trend of Fig. 20.

Page 45: The spread shown really is not very big (i.e., why do they say that the spread is significant?).

Figs. 21 and 22: Labels on all axes.

Page 55: The statement is made that the factor model parameters have no uncertainty by definition. Please explain. I have trouble with this statement. *They could have a more sophisticated set of factors (e.g., more than two, not all multiplicative), that would allow for more careful uncertainty assessment. They can make the statement that they are not examining the impact of having the current form on the results and the uncertainty assessment.*

Page 59… We await the results.

*As presented at the meeting: I was pleased with what I saw at the meeting, and rather than attempt to be exhaustive, and not complete all of the analyses, continue as planned to examine the most influential. They should concentrate on the uncertainties in the "differences" in exposures (this was presented, and one notes that the differences are small). They need to be ready to explain the skewed results, and a figure would help.*

**Comments on the Cox and Camlier memo:** First, thanks for doing this. I was disappointed that a bottom line recommendation was not given as how to treat the calculated uncertainty in terms of the precision of the standard. This would indicate that the precision of the standard should be to the nearest 1 ppb.

***Other:*** *As discussed with staff (Langstaff), a simple diagram of the rollback approach would help significantly.*

# Dr. Elizabeth A. (Lianne) Sheppard

## Review of the Ozone Staff Paper Second Draft

### Staff Paper Chapter 5

Response to charge questions – Overall comments

*Heath Risk Assessment (Second Draft Chapter 5 of the $O_3$ Staff Paper and draft Health Risk Assessment technical support document):*

1. To what extent are the assessment, interpretation, and presentation of the results of the revised exposure analysis as presented in Chapter 5 (and in the second draft Risk Assessment technical support document) technically sound, appropriately balanced, and clearly communicated?

I found generally the work in this chapter and the accompanying risk assessment to be well done, balanced, and reasonably communicated. There are a few places where I suggest improvements and I have included these in my detailed comments. Most important, the presentation and summarization of the clinical evidence for the $FEV_1$ decrement dose-response implies far greater certainty than I believe is warranted by the data. Figure 5-2 must be completely revised to reflect the uncertainty in the data. A new table summarizing the derived study results used for this figure should be added to the Staff Paper. The method for deriving the dose-response calculation, and the use of the logistic function must be re-evaluated. Furthermore the sensitivity of the results to the dose-response assumption must be assessed. The apparent certainty of the $FEV_1$ decrement dose-response has major implications, particularly since the estimated $O_3$-induced decreases in lung function are lower than were estimated in the last review. Because of the currently hidden uncertainties, this conclusion may not be correct and should be revisited after gaining new perspective on the certainty of the estimated dose-response function.

2. In general, is the set of health endpoints and concentration-response and exposure-response functions used in this risk assessment appropriate for this review?

Given the literature to date, the set of health endpoints are appropriate. I find it more difficult to infer causality from the epidemiological studies than the staff, but I agree that the endpoints and estimates used are reasonable choices given the current state of the science. It is important to state clearly in both Chapters 5 and 6 that there are many endpoints with good evidence of effects cited in Chapter 3 that are not incorporated into the risk assessment in Chapter 5.

3. Are the methods used to conduct the health risk assessment technically sound? Does the Panel have any comments on the methods used?

My main concern is the use of the 3-parameter logistic function for the derived lung function response results. Not only is the function problematic, but also the summarization of the data prior to selecting this function allows the perception of greater certainty than is warranted by the data to prevail. Because it has a large impact on the use of the clinical study estimates from the risk assessment, the approach to fitting this function and its sensitivity must be evaluated. See my detailed comments.

4. To what extent are the uncertainties associated with the health risk assessment clearly and appropriately characterized in both the second draft Chapter 5 and the second draft Health Risk Assessment technical support documents?

I was pleased with the effort to document the uncertainties. While it is difficult to know whether or not all uncertainties are well enough understood to even document, the lists appear thorough and it is clear that this issue was thoughtfully considered. In terms of risk communication, be careful that thoughtful scientific consideration of the uncertainties doesn't dilute the overall message.

Detailed comments

Please review all tables and figures (including captions) with an eye towards better labeling, particularly of figure axes and table columns. Always include units and enough details so the figure or table can stand alone out of context. For instance there are a number of figures that have the ozone concentrations/standards listed and also include "Recent (2004)" as one of the items. This is confusing out of context. It also may be helpful to give a detailed key of the first figure in a series.

Thorough cross-referencing with the technical Health Risk Assessment document would be helpful.

My concern is that focus on the risk assessment demotes in importance all other evidence in favor of $O_3$ health effects. Highlight a better in this chapter that the effects and endpoints used for the risk assessment are a small subset of the effects that were summarized in Chapter 3, as well as a subset of populations, locations, etc. An aspect of this is noted briefly on 5-7 9-12. It is too easy to focus only on the risk assessment results and overlook the other large body of evidence that is not included in the risk assessment. The write-up should help readers avoid doing this. This point also needs to be brought forward into Chapter 6.

Make sure percentages are clearly defined. Avoid whenever possible using number of events in the denominator. Instead use total population. If events are needed for some purposes, take care that counter-intuitive comparisons and results don't get highlighted.

5-4 lines 1-11: I'm not convinced that we can characterize the uncertainty of risk assessment in epidemiological studies as "considerably" greater due to location. I don't think we have great enough understanding of the sources of uncertainty to quantify its relative magnitude. Similar comment for the use of "considerably" in line 18.

5-5 and throughout: While observing consistent results across epidemiological studies may be a necessary condition for inferring a causal effect, I don't think that it is a sufficient condition. While it may only be semantic, I find the judgments about causality as stated in this document, particularly with respect to the epidemiological studies, difficult to accept. One reason is that the epidemiological study relative risk estimates come from the time series study design and we have limited understanding of the effect of this study design, particularly for such a highly reactive pollutant such as ozone that does not penetrate well indoors. However, based on the perspective of the precautionary principle, I do not object to the selection of endpoints and their use in the risk assessment.

5-12 and throughout: Notation in this document is at times confusing. Anything staff can do to minimize the variants of notation through the whole document, clarify the definitions, and unify the presentation, would be helpful. One suggestion is to prepare a glossary of symbols for Chapter 5 and/or for the TSD. (Note: eventually I found a glossary for some terms in Appendix B.2. However *B* is not defined there. At a minimum refer readers to this appendix in Chapter 5.) In the process of preparing the glossary, instances of redundant and confusing notation will be more readily identified and can then be corrected.

5-14: It is unclear to me why the categories for "as is" and "background" would ever be different (i.e. why distinguish between *i* and *j*?). Is it possible to simplify the equations from the most general forms and still capture all of the work that was done? Or are there examples where this isn't possible?

5-16: I have difficulty with the use of the 3-parameter logistic function for estimating the dose-response curve. While from the existing figure it appears to be better than the previously used linear dose-response curve, it is not apparent to me how sensitive the results are to this assumption and how necessary it is to make a parametric assumption, particularly from a 3-parameter model fit using only 5 data points. The biggest problem with the logistic function is that it is parametric so that a small change anywhere on the curve may have a large impact in the region of primary interest – at the low ozone concentrations. In particular, note that the threshold response rate is estimated from these data although there are no study results at the higher exposures to verify this estimate. (As an additional point, I don't recommend excluding data that could be informative for fitting this model and thus suggest the dataset be expanded to include studies conducted at higher concentrations – e.g., the Follinsbee studies cited in the 1996 CD, particularly if the logistic function will continue to be used. Take heed from the O-ring data analysis on a restricted dataset used to justify the decision to launch prior to the Challenger disaster.) The logistic function is by definition symmetric, so small changes to the threshold estimate could have a large impact on the shape of the curve at low concentrations and consequently on the risk assessment estimates that are so heavily weighted towards the smaller concentrations. I have several follow-up suggestions: 1. Consider using a more non-parametric approach to smoothing as the main analysis, and certainly at a minimum as a sensitivity analysis. I suggest considering a natural cubic spline with one knot or pure linear interpolation. 2. Consider weighting the contributions of each data point by their uncertainties. Better yet, start from the raw data, not data averaged at each exposure level. 3. Particularly if a parametric function is to be used, include additional studies to better pin down the height of the curve. 4.

Conduct a sensitivity analysis for a few different curves, including a linear function if indeed that function is consistent with the data when the uncertainties are included. If the risk assessment isn't particularly sensitive to changes in the approach to smoothing, then it will be less important to focus additional attention on the approach to smoothing at this late date. Based on the comparison of these new results to the previous review, the risk assessment is very sensitive to the shape of the selected dose-response function. Therefore more needs to be done to address that sensitivity.

Figure 5-2 (and 3-3): Completely revise this figure. Show study-specific data, not summarized data. Include uncertainty estimates on the figures. Show the point estimates and uncertainties for all the studies (by e.g. jittering around the target concentrations). Also add a table that gives all the results used to create the data in these figures (or cite the appropriate table in the CD). Chapter 3 (Appendix Table C-2) has the data, but it is not in the same form as the data used in this figure. There are no counts given in the table. Thus a new table needs to be developed and included in Chapter 5 to show the data used in the figure.

5-18: The "Bayesian" approach appears to be only Bayesian in flavor. For one thing the "data" in the "Bayesian" analysis are not data, but predictions from the logistic curves. While it is a commendable effort to quantify uncertainty in these dose-response curves, there is no recognition of the uncertainty due to the selection of the function itself. Also the data are not used directly.

5-22 13-23: This is a place where including the notation along with the words will help the reader immensely. I'm still not sure I'm completely straight in this section as to when "baseline incidence rate" refers to $B$ (the more typical use outside of this document) and when it refers to $y$. On first reading it was only on page 5-29 line 30 that I finally began to understand that $y$ is the baseline incidence.

5-24 17-30: Drop all this and just give the definition of notation, i.e. $\exp(\beta\Delta x)=RR_{\Delta x}$.

5-35 16-18: I suggest starting that sentence with "One" rather than "The". I think this is one possible explanation but not the only possible explanation as is currently implied.

5-49 21-22: The reason to use the highest 8-hour monitor when comparing to the 24-hour average is not clear. Epidemiological time series studies with 8-hour average ozone as the exposure do not use the highest monitor over a spatial region. These appendix plots are examples of plots that need to be more clearly labeled and described in the caption. Also the distribution of the maximum is related to the total number of monitors. Include this information (the number of possible monitors) in each plot.

5-56 21-23: It is helpful to know that the risk estimates are based on the accumulation of events over seasons that are different lengths. This means from one perspective, comparisons across cities are "unfair". I think this information belongs in the table as a footnote or extra column (summarizing the percent of year for the ozone season) as well. Another idea is to add an additional table in the TSD to allow cross-city comparisons for the single highest average

exposure month (or some other reasonable choice that would make the estimates from all cities temporally comparable).

5-71 9-11 and 31-34: This comment elevates the importance of my concerns regarding choice of the logistic function.

5-71 15: Figure 5-12 missing. Clarification: This is a typo; should be Figure 5-2.

5-73 8-10: It is unclear to me why New York was singled out for discussion.

5-76: I think there are also uncertainties due to the features/behavior of particular study designs, particularly the time series study design that relies on aggregate data and does not incorporate a one-to-one link between exposure and outcome on individuals.

Minor comments:

5-12 3: Why is there an *i* subscript included for $FEV_1$?

5-24 13: Replace this line by: subtraction equation (5-3) from equation (5-2) and multiplying by y/y which yields:

5-33: An asterisk appears in the table without a matching footnote.

5-36 and throughout: I would prefer the term "sampling variability" over "sample size considerations".

5.57 16, 17; 5-73 22: Periods missing.

5-76 43: Replace line with: model uncertainty (i.e., uncertainty about the selected model, e.g., the shape and magnitude of the

5-78 25-27: More information is needed to easily access this paper.

## Staff Paper Chapter 6

General comments:

I found this chapter less well done than chapter 5. I found several errors, suggesting to me this chapter contains more than an acceptable number of errors. I felt that the composite body of evidence of $O_3$ effects summarized in chapter 3 was discounted, perhaps because many of the results cited in Chapter 3 aren't easily quantified in a risk assessment. The force of scientific evidence should not be lost with focus on a limited risk assessment. In contrast, a few studies (e.g. the Peters et al 2001 pilot study results) were given inappropriate attention relative to other evidence in the literature.

Most important, I did not find the evaluation of the risk assessment well grounded in public health considerations. I think the first order of business for this chapter is to clearly define and justify the rationale for criteria for public health benefits of any standard. I suggest public health significance should be based on absolute rather than relative criteria. While the risk assessment is admittedly restricted to a subset of endpoints, population groups, and geographic locations, numbers from the risk assessment can be used for quantifiable estimates of public health benefits.

The second problem with the focus in this chapter is that it gives relatively too much weight to the necessarily restricted risk assessment (restricted to specific populations, endpoints, locations) compared to the full body of scientific evidence for $O_3$ health effects summarized in Chapter 3. This problem starts in Chapter 5 and is carried into Chapter 6.

One of the effects of the focus of this chapter is that the conclusion of possibly maintaining the current standard was presented as defensible from some points of view. I think the body of evidence only supports lowering the current standard. The chapter should be revised to focus on a vision for acceptable public health impact. Then all the evidence should be brought to bear on this vision without overweighting those results that have an accompanying risk assessment. This will produce different conclusions.

I found the reliance on relative comparisons using cases as the denominators misleading (e.g. see Table 6-3; also Figures 6-1 through 6-6). Percents based on number of cases are sensitive to variation in the number of cases from year to year and this creates apparent differences between percentages that are not real.

- For instance for incidence in Table 6-3 (see also 6-21 lines 17-18) the incidence differences for recent vs. just meets the standard are identical but the percentages make it appear that it is better to reduce $O_3$ in a clean than a dirty year when in fact the absolute improvement is identical in both years.
- The figures show percent changes relative to the current standard.
    - It does have value in normalizing the data and allowing comparisons between locations. However, since many locations already don't meet the current standard, this point of reference can be questioned. At a minimum keep the point of reference the same but also include the actual data as an additional exposure on the figure.
    - Percents can also mask the public health significance by magnifying small absolute differences and the converse. For instance, going back to Table 3-11 in the TSD and comparing FEV decrements in Atlanta, note that the **relative** change for 0.084/4 to 0.80/4 is 11% for FEV greater than 10% and 24% for FEV greater than 15%. However, for FEV greater than 15% affects 11,000 occurrences while for FEV greater than 10% affects 65,000 occurrences. This is so much larger in **absolute** magnitude that it cannot be ignored. See page 6-33 lines 13-15 for an example of where the relative comparison ignores the large differences in absolute effects.

Add some discussion in this chapter about atmospheric chemistry and the importance of reducing oxidants on reducing overall pollution. Incorporate recognition that O3 is an indicator.

Review the language used in this chapter for clarity.

Details:

6-10 30: The $O_3$ data used in Sheppard 2003 and the original 1999 paper are predominantly for the warm season. Only one year of 8 years of data included winter measurements. Also correct the reference.

6-11 8-10: This conclusion is confusing. Based on the 98$^{th}$ percentile ppb averages, clearly most or all of the monitors in the Burnett et al 1997b paper met the standard.

6-11 29-31: Since the epidemiological studies estimate a linear dose-response relationship, and most when they check this assumption find that it is well supported by the data, does it really make a difference whether the areas in which the studies were done met the standard? Is this an opportunity to incorporate the concept of tolerance?

6-12 22-23: Can this be stated more clearly? Define the relevant air quality statistics.

6-12 27: Revise/add to the sentence as follows: "…likely to produce different estimates for reasons such as seasonal differences in pollution or confounding…"

6-15, Table 6-1: Here is an example of where the percentages are defined and their use is helpful for the interpretation.

6-17 15: Drop "for"

6-18 Table cell "just meeting" and "occurrences for FEV >20" the percent should be 90% not 85%.

6-21 17-18: Note that the absolute changes here are identical while the relative changes are different! From a public health point of view, there is no difference between the reductions in the two years.

6-22 17-18: Note that the estimate from Bell et al is for all year even though the exposure data are restricted to the $O_3$ season.

6-22 Table 6-3: Does it make sense the 2002 incidence per 100,000 is higher than the recent air quality incidence rate?

6-28 11-12: Improve wording, e.g. drop the second "important".

6-30 4: Citing only the pilot study results ignores the main study results which should be viewed as scientifically better grounded.

6-32 1-3: This sentence is unclear. Would this argument benefit from the evidence in the most recent Bell et al paper (2006) that shows the same effect estimate as days are excluded at progressively lower exposure cut-offs.

6-32 20-27: Here is an example where the relative values may not give a full perspective.

6-33 12-15: As noted above, the relative comparisons hide some important information. At a minimum make it clear these comparisons are relative by e.g. adding "relative" after "greater" in line 13. Better yet, revise to acknowledge the absolute impacts as well. Incorporating absolute changes will also force the discussion to be more mindful of the very limited subsets on which the risk assessments were done.

6-33 29: Define reduction. Reduction in what? This is an example of a more pervasive issue in this chapter.

6-42 13: Figures 6-1 and 6-2 don't have a and b.

6-49 36: Line missing.

## **Staff Paper Chapter 3**

Specific comments:

Exposure error section 3.4.2.1, particularly page 3-37: This page needs improvement to convey a clear message. As a preamble, it can be helpful conceptually to partition total personal exposure into ambient and non-ambient sources. The ambient source can be simply assumed to be a product of ambient concentration and attenuation due to the building filter (call this $\alpha$). Absent residual measurement error and confounding, an individual study such as a panel study using personal exposure to a pollutant will estimate toxicity of the pollutant (e.g. $\beta$). Assuming measurable exposure, a panel study using ambient source exposure, will also estimate toxicity ($\beta$). The same panel study using ambient concentration instead of ambient source exposure will estimate attenuated toxicity (toxicity times $\alpha=\beta\alpha$). Similarly, a time series study using ambient concentration instead of personal exposure will also estimate attenuated toxicity ($\beta\alpha$). This is the point being made in Sheppard et al (2005). Lines 5-9 don't clearly summarize Sheppard et al.'s simulation study results. A time series study does not estimate the same effect as an individual study with personal exposure would estimate. Sheppard et al (2005) showed no further attenuation as one might anticipate be present due to the poor correlation between personal exposure and ambient concentration. Their results showed that when total personal exposure was the true exposure, a time series study estimates the effect of concentration, which is the product of toxicity of the exposure and attenuation due to the building filter.

3-37 24-25: I note that point (2) from the Zeger et al paper also results in the point made above (i.e. that estimates of $\beta$ aren't the same as estimates of $\alpha\beta$). The way this page is written it appears that these two points conflict.

3-38 10: The Brauer et al. reference is missing.

3-39 lines 1-4: Drop the rest of the sentence after the comma. Not only does this confuse the reader, but it is irrelevant to the Staff Paper since the epidemiological risk assessments always use concentrations as the exposure. Line 5: change "these" to "the epidemiologic study". It is correct that the health effect parameter in air pollution time series studies doesn't estimate toxicity (called potency in line 5). However, the relative risks for increments of concentration, while not directly interpretable as potency of $O_3$, are interpretable.

3-43 2: This Peters pilot study result should not be cited over the primary study result.

AX6-4: The >6 hr exposure results are not summarized in a way that is used later in the SP. This should be rectified here or in a later table in Chapter 5.

## General Staff Paper Comments

Where possible, enhance the document with comments to help the reader navigate and look forward and backward in the document. Add cross-references whenever possible. To look forward, mention planned uses of calculations. To look backward, mention source numbers from previous chapters. Make sure every figure and table has a thorough caption and good labeling so it can stand alone out of context as much as is feasible. Most readers will not digest this document from start to finish. Therefore anything that can be done to help a reader starting in the middle find relevant information and assumptions from other parts of the document will be very helpful.

## Health Risk Assessment TSD

Comments that weren't completely covered above in SP Chapter 5 comments:

3-4, section 3.1.3: This section does not adequately document the details used for the calculations. **The raw data must be presented**. If the averages for each exposure are to be used for deriving the dose-response function, then the calculations and intermediate steps taken to obtain those averages must be presented. (I would prefer that the analysis be redone using only the raw data presented as estimated probabilities of response at the different exposures.)

3-4: Document that the logistic function fits clearly better than a linear function once the uncertainty in the data is included. Once the uncertainties in the study results are incorporated, I suspect this conclusion is not correct. If the logistic function is to be retained, extending the plots (Figure 3-3) to include higher exposures in order to show the threshold level would help with physical interpretation of the dose-response function.
4-40: I agree the lack of daily incidence data shouldn't be a problem, but can't you use the city-specific mortality or admissions data comparable to those used in the epi studies?

**References**: Here is one study I found from a web search that includes >6 hr exposure above 0.12 ppm. I understand from my clinical colleagues that such studies were reviewed in the previous CD.

Kehrl HR, Peden DB, Ball B, Folinsbee LJ, Horstman D. Increased specific airway reactivity of persons with mild allergic asthma after 7.6 hours of exposure to 0.16 ppm ozone. J Allergy Clin Immunol. 1999 Dec;104(6):1198-204.

# Dr. Frank Speizer

Review of second draft of staff paper on Ozone Dated July 2006
Submitted by: Frank E. Speizer, CASAC

Chapter 1

Notably on page 1-3, line 25-30 the reading of Section 109(d)(2) states that an independent committee …. shall recommend to the Administrator and new ….or change in the standard. That committee since the early 1980s has been CASAC and it therefore seems to be more than an advisory committee that only approves or rejects what EPA staff does.

Chapter 2

Page 2-5, line 17-19, in describing the findings in table 2.1 for NOx emissions it is clear that emissions have come down, in the face of increasing cars on the roads, and it probably worth indicating the reasons,, e.g. Better controls, catalytic converters, fuel processing, change in reference methods??? Say something. Same applies to table 2.2.

The box plots in Figures 2.2-2.4 of the 1, 8 and 24 hours show striking uniformity across the country. This is indicated by suggestion in the text that the expected excesses in California are not localized to California alone, but it might be more explicit to state that the average levels are similar and for the 8 hour the levels above .07 represent 10% of the levels recorded in all parts of the country. (I note this is roughly seen in Figure 2.6).

Text to page 2.40, mostly describes what is in figures, without providing much in the way of interpretation. If one just focuses on the 8 hour data it seems pretty clear that in many of the cities, and for many of the time trends there are a substantial number of excedences each year ( e.g. Figure 2.17 shows monthly averages of 8 hr $4^{th}$ largest over .08 ppm with no particular improvement with temporal trend. This is not discussed (maybe it is discussed elsewhere). Even if it is, it should be mentioned here.

Page 2.48, last sentence. Would it not be worth a statement that each of the cities have remarkably consistent finding for both month and time of day? Again this might be discussed further in later chapters, but certainly no reason not to say it here since the subject is raised.

With regard to Martin Questions:
See above. Although it appears that the data are presented and brief descriptive analyses are presented in the tables and figures, some interpretation needs to be added to text.

Chapter 3

The set of appendix tables (Appendix 3B and 3C) are excellent. They are laid out in a logical fashion, and should be a model for the future.

Page 3.6, lines 7-12. The terms triangular exposure and square wave exposure represent jargon of the field, and may not be understood by most readers. Suggest definitions be supplied.

Page 3.9, line 27-28: For clarity the odds ratio of 1.37 should be indicated in this sentence to be per 30ppb (if it is).

Page 3.12, line 7-8. Need to reconcile this statement with the first paragraph of section 3.3.1.1.3 on same page. It seems inconceivable that inflammation is not involved in airways responsiveness and is involved in exposure in the lung. Certainly some part of the airway system, albeit smaller airways, which contribute less to responsiveness, are involved in the inflammatory response of the lung. In fact rest of page and next goes on to indicate that some inflammation must be present in the airways.

Section 3.4 starting on page 3.34-4.47, title is a little confusing. Certainly what has come before is much of the epidemiological as well as other evidence and I would have called it an assessment thereof. To label this section "Assessment of …" seems odd. After reading it I come away with the feeling that it is an attempt to justify what has come before on the basis of Hill's postulates. That's ok, and maybe a little more thought to make the title of the section more appropriate is in order. There is also a style issue in this section. Each of the subsections concludes with essentially a quote from the CD. This really isn't an interpretive statement but more a restating of what the CD states.

Page 3.62, line 14: Suggest change for clarity to: … 2 fold greater decline in…

Martin Questions

1. The evidence is presented in a sound effective manner. In fact it is not clear to me what will be different in the subsequent chapters on health as much of the information that will be contained can only be what is here in greater detail.
2. The discussion of key issue is also well presented. It is organized effectively, however, if anything toxicology and controlled human exposure studies seem to be played down somewhat. Other epi issues are well discussed.
3. Threshold is adequately discussed.

Chapter 4

The data in table 4.4 suggest that a relatively small number of diary-days, particularly for the children aged 5-18, were actually obtained. In addition, the range of days varies from 42 to 4332. Yet, as far as I can tell when these data are used each city is given equal weight. When these data are used shouldn't the results be weighted by something like an inverse square term to take into account the relative uncertainty of the data? Maybe this is done in the generation of values but it needs to be specified somewhere.

There is mislabeling of the Figures in the text (Figure 4.1 and Figure 4.2 in text on page 4.30, line8 instead of 4.2 and 4.3. This continues as a problem on line 16.

Martin questions

I found this chapter relatively easy to read and quite informative in its use of the other related documents from which it draws its information. However, I have a problem in that much

of the information used is not from the CD but from the other documents, some of which are still called drafts. I know that we recommended the use of "grey literature" but I would not consider the ancillary documents here that grey literature. This is an issue I think we need to discuss as a group and choose and go on record as endorsing the approach. I think it is ok but I think CASAC needs to sign on.

The methods used are reasonably sound. Uncertainty is referred to, in citing another document, but probably needs to be explored further. For example there are no error bars on any of the later figures that give estimates of population exposed to specific levels. I do not think these must be calculated for each point but some general examples need to be given. Because there are such differences in the data by each city, are there ways they can be combined that would allow for more robust assessments, perhaps by using dummy variables for city while considering all data together. Other more sophisticated statistical approaches might be considered.

Chapter 5

Page 5.7, lines 10-12: This is an important caveat. The suggestion is that the current quantitative risk assessment gives only a minimum estimate of true short term risk. This needs to be discussed further.

Page 5.8, Para beginning line 8: This suggests that by not using emergency room visits and school absences in current risk assessment, the risk assessment is leaving out large potentially important health as well as economic impact of exposures.

Page 5.20, Figure 5.3a, b, c: In contrast to Figure 5.2, each of these components extend the curves beyond the data (0.12ppm) and seem to imply the exposure-response curve is approaching an asymptote. I do not think this is justified and suggest curves be truncated at 0.12ppm.

Martin questions

Assessment seems technically sound and methods seem appropriate. The question that comes up is to what degree is the concern about uncertainty a reflection of the lack of data on endpoints that are in fact not used in the assessment. E.g. all might agree that cardiovascular end points lack specificity and lead to uncertainty but in fact they are not used in the calculations of the assessment. With regard to the technology used, to the degree it is dependent on the Draft Risk Assessment technical support document, I would think that a critical evaluation of that report and its final status will need to be documented somewhere in the process before it can officially be used as a basis of the assessment.

Another point of concern is the details on the suggestion that the effective mortality is related to tolerance. The issue is raised but not adequately discussed.

Chapter 6

The chapter is reasonably well written and essentially summarizes what comes before. One issue mentioned below seems to me not adequately addressed.

Page 6.6, sentence beginning line 23. This needs rephrasing or could simply be left out. One could just as easily argue that the no threshold was found as to neither supports or refutes, as it is all a matter of the level of exposure used in measures of effects. The rest of the paragraph discusses this issue adequately.

Page 6.17, line 6, typo: 2004

Page 6.30, end of paragraph at line 13. Two issues that are not adequately mentioned in this chapter but might be important modifiers in a discussion of the averaging time 1) are the quenching of O3 by NOx, that may be occurring in heavily traffic situations and 2) the natural night time drop of O3 that occurs as sea level sites (thus making the 24 hour measure an underestimate of high exposure for shorter periods). Surely these are part of the logic of selection of 8 hours as the measure of choice.

6.47, line 16, typo: microenvironments

Martin questions
The combination of evidence-based and quantitative exposure-and risk-based consideration seems appropriate, but as mentioned in discussion of chapter 3, much of the risk assessment is based on draft documents that have not had adequate peer review. I know that a request has been made to have CASAC review (and some of us will clearly get to that) but I am not sure that will be adequate and we will need to discuss the issue.

The presentation on Page 6.44 on whether to retain or change the current standard is appropriate but as presented results in an unbalanced form. Paragraph **a** indicates the standard could be retained but bases this decision on a very narrow view of the data. This could be interpreted as saying that only controlled human study data with a 15% decrement in lung function is an adverse effect that should be considered. (The paragraph goes on to say that this would ignore much other evidence, which the administrator could choose to do). However, there must be a way to indicate some of the relevance of the other disciplines that need to be considered in making a decision to retain the existing standard. Whereas paragraph **b** is more thorough in considering the data, and emphasizes the potential importance of albeit less precise but perhaps as relevant, population based data.

One way out of this dilemma would be to reverse the order of the presentation and give the alternatives for changing the standard before the idea of retaining the standard. (I recognize that this would be different than the way this is done in the past, but the logic of not using all of the data seems inappropriate otherwise).

With regard to the key uncertainties and research recommendations it is disappointing to realize that we have made so little progress since 1997, although I am not so sure I would totally agree

in that conclusion. In fact, because of the work that has been done, particularly in working out the methodologies for PM, a lot of the methods questions have been answered. New issues have arisen, and these have been summarized reasonably well. Perhaps more could be made of the approaches that might be used to explore mechanisms and potential gene-environmental interactions that may be important in identifying subgroups of susceptible populations.

---

Review of Analysis of Uncertainty in O3 population exposure modeling
Dated July 24, 2006
Submitted by Frank Speizer, 8/23/06

General Comment:

This document claims that in its next iteration it be included in its final form to inform the final Ozone Staff Paper. It seems clear that unless a substantially revised document already exists (and this one claims it is waiting for in put from CASAC and the public) that this will be an impossible task. Much of the document is a promissory note of what the author is planning to do. Simply reviewing the existing literature cannot be accomplished without a Herculean task to have some thing ready in the next month that will be informative for the final staff paper. This is another example of too little too late and my suggestion would be to put this aside and find something more useful to do with the remaining time before the court ordered deadline is upon us. I certainly would not be able to sign off on this document and certainly there will be no time to review the next iteration.

Specific comments:
Page 5, para 4, line 5: Recognizing that the author is referring to another document, it is not clear where these SD are coming from. Need to indicate that these are simulated models with arbitrary selection (if that is what they are).

Page 8, first para: Editing: This says the same thing with the same words as that on the top of page 4

Page 11, para 2, end: If 2003 and 2004 are not correlated, then what is the meaning of indicating the averages from the table (I assume author means all the numbers) have average values as given? Further what does it mean that the range of bias for 2003 is -7.35 to +11.28!?!

Page 13, first paragraph. It is not clear how these statements will be used. In fact meaning is not clear.

From here on out the promissory nature of the discussion becomes increasing frustrating and indicates to me that the tasks suggested will never be completed in association with the time frame proposed.
Page 15, end of first paragraph. What is the use of indication of what is said and is it logical to assume that the interpolation will be unbiased?

Next paragraph:  We plan to perform…It hasn't been done, will it be done for this report or future reports?
Figures 6-8:  Should there be error bars on these curves?

Page 19, Table 8:  Not clear why all upper bounds are equal to 1.0

Figure 10, error bars?
Page 20, Table 9:  Does this make sense that all values save in-vehicle, interstates are virtually the same?

Bottom of page 20, top of page 21, first 2 para.  Too much work is being planned for an October deadline.

Because the remainder of the document suggests that work will be done, I see no way to integrate this into the staff paper at this time.

# Dr. James Ultman

Chapter 6 in the OAQPS Staff Paper for Ozone (2nd draft)
Reviewer: JS Ultman
Date: 8-25-06 (revised)

This chapter is well-written and logically presented.

Based on the scientific evidence, exposure analysis and risk assessment, I conclude that the current 0.08/4 standard does not adequately protect asthmatic children from respiratory adversity due to ozone exposure. In addition to exposures above the "level-of-concern" and adverse FEV decrements predicted to occur when just-meeting the current standard, a number of independent studies indicate a significant increase in emergency department visits and hospitalizations due to the exacerbation of asthma (fig. 3-4, pg 3-53).

Inflammatory processes are an important marker of tissue remodeling and potential long-term lung injury due to ozone inhalation. The use of 0.08 ppm as a "level-of-concern" in this chapter is consistent with the ozone concentration at which inflammatory mediators have so-far been observed in human studies. Clearly the selection of this level-of-concern has an important effect on the number of individuals and the number of accidences per individual that is predicted by the exposure-activity model, and therefore the rationale of using 0.08 ppm should be better explained in this chapter.

The chapter summary states that the quantitative risk assessment is now appreciably lower than in the last review (pg. 6-44, lines 13-16). This conclusion is based on an exposure-response curve for $FEV_1$ that is my view has a great deal of uncertainty. Thus, this statement is misleading and should be highly qualified, or entirely removed from the chapter.

Some other comments:

Page, line

| | |
|---|---|
| 6-4, 10 | Delete first "human |
| 6-7, 24 | Delete "also" |
| 6-13, 35 | It would be useful to give your rationale for choosing these subgroups of the population for analysis. |
| 6-22 | In table 6-3 and tables that follow, it is not clear what incidence means in the context of "incidence per 100,000" and "percent of total incidence." |
| 6-27,10-15 | What are the "less serious" lung function effects you referring to? Why are they "less serious?" |
| 6-40, 18-21 | This sentence asserts that there is a 90-100% reduction in exposures-of-concern estimated at the 0.07/4 standard. Are these percent reductions obtained by rolling back the actual air quality for 2002 and/or 2004 to the 0.07/4 standard. If so, it appears from table 6-1 that similar reductions are possible using the 0.08/4 standard. I find this hard to believe. |

# Dr. Sverre Vedal

**August 2006**
**Critique of 2nd Ozone Staff Paper draft (chapters 3-6)**
**Sverre Vedal**

**Chapter 3 (Health effects evidence)**
While in general this chapter is a credible basis for the risk analysis that follows, there are inconsistencies and some inaccuracies, some of which are due to the necessary reliance on a too hastily prepared CD. There is appropriate use of cautions when the data are not as strong as they might be, but this is inconsistent. A few points:

1. Measurement error discussion is convoluted and confusing, and contains some mistakes.
The primary issue in the use of central ambient monitors for ozone in time series epidemiological studies is whether they provide any information at all in reflecting daily personal ozone exposure in the susceptible populations, especially that of the debilitated elderly in the case of the mortality findings. The evidence on this issue is split, but as opposed to what is claimed in this chapter (p. 3-38, lines 26-29), there is information on this. The two Sarnat papers, one from Baltimore and the other from Boston, assessed relationships between personal and central ambient concentrations in elderly subjects with COPD.

The discussion on the impact of various types of exposure measurement error is incorrect (p. 3-37, lines 22-30) or confusing. The difference between true and measured ambient concentrations is an example of classical measurement error and hence results in bias of effect estimates to the null, not just an increase in standard error. Further, claiming that the difference between average personal exposure and ambient concentrations results in "attenuation of risk" (line 29) is confusing (also p. 3-38, line 2 and 8, and p. 3-39, line 4 and 18). This statement is generally used when, due to classical exposure measurement error, the effect estimate is biased downwards, and hence attenuated or underestimating true risk. The effect of ambient concentrations themselves is underestimated. Here, as in the case of PM, the effect estimate is estimating effects of changes in ambient concentrations. But, what is being referred to here is the difference between ambient and personal concentrations, with personal concentrations often being substantially lower. In this case, the effect of ambient concentrations is not being attenuated, at least not in the same sense. They are only attenuated if one is trying to estimate the effect of personal exposures for concentrations that are similar to ambient concentrations, which is in fact not the case. This is not, therefore, a reason why we might be "underestimating true public health risk" (p. 3-39, line18).

The assessment of plausibility still does not adequately discuss the issue of ozone exposure in the mortality time series studies (referred to on p. 3-35, but does not come up again). It is again claimed that central monitoring concentrations "may serve as adequate surrogate measures for mean personal exposures experienced by the population" (p. 3-38, lines 21-22). How can this be if in some settings there is no relationship between individual personal concentrations and central concentrations? If <u>mean</u> personal concentrations correlate with central concentrations, and individual personal concentrations do not, then I presume some sort of ecologic error is at play here. I dispute the conclusion that "the use of routinely monitored ambient ozone concentrations as a surrogate for personal exposures is not generally expected to change the principal

conclusions from ozone epidemiological studies" (p.3-39, line 10-11). On the contrary, we have little insight as to what we would find had we actual exposure measurements.

The argument that central monitoring sites provide more valid measures of personal exposure in respiratory admission time series studies, specifically, because most such admissions are due to asthma admissions, most asthmatics are children and children spend more time outdoors, is especially convoluted and unconvincing.

2. Cardiovascular mechanisms and effects.

The discussion on plausibility of potential mechanisms underlying cardiovascular effects should begin with an acknowledgement that we don't know much about these. The current presentation (p. 3-55) comes off as a hodgepodge of "the usual suspects," some of which are a real stretch and serve largely to undermine credibility of the more realistic possibilities. The discussion on the role of ozone in PM-related cardiovascular effects is hypothetical (note the ultrafine discussion, in particular), yet are presented as though these interactions are actually playing a role (p.3-59, lines 4-7).

3. Low concentration ozone effects on lung function.

First, Adams (2002, 2006) concludes that 0.04 and 0.06 ppm responses were no different than filtered air responses. It is necessary to make it clear (especially in Ch. 6) that the estimated effects in Fig. 3-1 (p. 3-7) are derived from Adams' data, but are not in fact reported in his paper. It now becomes clear that the percent of the relevant population estimated to experience FEV1 declines at these low concentrations, estimates that are central to the risk analyses in chapters 5 and 6, are based on a total of the 30 subjects in Adams' papers (i.e., 3 subjects for 10%). This area of low concentration responses is an area that clearly needs more research.

Because these low-level effects have a very prominent role in the risk analysis in chapters 5 and 6, the summary table (p. 3-78, table 3-4) should include them in some fashion.

4. Chronic effects on lung function: overstatements and inconsistencies.

The conclusion on p. 3-20 (lines 16 and 32) is correct: we have "little evidence" for these. Why then do we get statements such as, "these long-term exposures may be related to changes in lung function in children" (p. 3-56, line 30), "reduced lung function development in children which have been observed in epidemiologic studies" (p.5-56, line 25), and "recent studies of lung function changes in children living in cities with high ozone levels support the conclusion that long-term ozone exposure may play a role in causing irreversible lung damage" (p. 3-84, lines 6-8)?

5. Coherence of respiratory effects.

It is repeatedly claimed that the lack of consistently seen in the increased effect estimates for COPD mortality may be due to inadequate power (e.g., p. 3-29 [line16], p.3-52 [line 18]). Note that this is in stark contrast to the case for PM where such relatively larger effects are almost always seen; and the data used for the PM analyses have no better power than these for ozone. Also note that the large meta-analysis (Bell 2005) found smaller effect estimates for respiratory causes than for total mortality (p. 3-30). This has implications for the argument for plausibility based on coherence of the respiratory findings. Parenthetically, why are not the studies used in the Bell meta-analysis (2005) included in the respiratory mortality studies in Figure 3-4 (p.3-53)?

Smaller or detailed points:
Page
3-7 (Figures 3-1A→C). These do not indicate "the portion of subjects tested having FEV1 decrements in excess of 10%," but rather give the distributions of % decrement. Also, clearly the findings are somewhat dependent on the subjects included, since the distribution of decrements in the McDonnell subjects for filtered air is about the same as that in the Adams subjects for 0.04 ppm ozone.
3-24 (line 12) and 3-79 (line18). Note that Gong did in fact find an effect on increased heart rate, as stated correctly on p.
3-25 Correct the referencing in the last paragraph – the extension to NMMAPS analyses is only in the Bell 2004 reference
3-26 Regarding adequate control for temperature, control for single-day effects or matching on a day (as in the case-crossover approach, lines 28-31) is probably not an adequate assessment of the adequacy of temperature control.
3-27 Effects using 1-hr and 8-hr averages in NMMAPS are not "distinctly" larger (line 13) than those for 24-hr averages; "slightly" larger would be more descriptive.
3-32 The ozone effect in the ACS II study are not always nil (line 18). For one of the two ozone monitoring periods, the effect was clearly positive.
3-35 (line 18) There is an interesting allusion to the utility of time series studies here which doesn't seem to come up subsequently in the SP.
3-38 Reference to Brauer 2002 (line10) is needed in the reference list, if it has not already been added.
3-44 (line 6) Note that Bell (2006) found no difference in estimated effect even when all days with 24-hr ozone concentrations < 20 ppb were excluded.
3-45 (lines 3-5) Note that the evidence of no clear threshold comes largely from epidemiological studies, and partly from clinical studies, but not from toxicological studies.
3-56 (lines 26-27) Another possibility, of course, is that in fact there are no effects on humans of long-term exposure.
3-61 (line25) These (5% decrease per 40 ppb increase) are not *large* decrements in lung function.

**Chapter 4 (Human exposure)**
1. It would be helpful to have the estimated exposures for current (2002 & 2004) levels displayed in Tables 4-8 & 4-9 (p. 4-32) and Figures 4-4 → 4-21 (pp. 4-33→4-41), in addition to only those for just meeting the current standard and alternative more stringent standards. This would be analogous to the way estimated effects are displayed in Ch.5 (Figures 5-5 → 5-9 [pp.5-58 → 5-65]).

Small points:
Page
4-30 (line 8 and 16) Correct the figure reference numbers.
4-31 (figure 4-3) It is still not entirely clear why the estimated exposures are so low in Los Angeles.

**Chapter 5 (Risk analysis)**

1. Multi-city mortality estimates vs. single-city hospitalization estimates.

Note that estimated multi-city mortality effects are qualitatively different from hospitalization effects, which are largely single city estimates (p. 5-28, table 5-1). Not only is there the well known phenomenon of publication bias to contend with in using single-city estimates, but, the Bell (2004) estimates for a single city are Bayesian shrunk estimates which are not the same as crude estimates of single-city effects, even from a multi-city study. Therefore, in using the mortality and hospitalization effect estimates in a parallel manner, it should be made clear that these are in fact qualitatively (and therefore quantitatively) different.

2. Confusion on the effect of exposure measurement error.

Because ambient concentrations are two- to four-fold higher than personal exposures does not mean risk estimates are attenuated (p. 5-36) (see my comments on this point for Ch. 3 [point 1]). Yes, potency is underestimated, but underestimation of risk is due to classical measurement error, not this type of differential measurement error. Risk would only be underestimated if exposure were to equivalent ambient concentrations of correct potency, which is not the case. The confusion occurs in making use of the analogy with PM.

3. Striking findings.

It is striking how much all of the estimated health impacts (from lung function decline to mortality) are due to exposures at concentrations below 80 ppb. Obviously this is due to the predominance of days with these concentrations.

4. Relationship between 24-hr average concentrations and those from the monitor with highest maximum 8-hr average (p.5-73 [lines 38-40] and Appendix 5A.2).

I'm not sure what point is being made here. While it is true that these are about half as high, it is likely that 8-hr and 1-hr averages are not nearly so different from the 24-hr average. Of more importance, the estimates of effect from the mortality time series studies using these different ozone metrics are not much different, as demonstrated in the Bell study (2004).

Small points:

Page

5-32 (table5-3) Note that "respiratory" as a single category is defined differently from "respiratory" in "cardio-respiratory"; also, this restricted definition of "cardiovascular" (limited number of ICD-9 categories) probably explains the relatively smaller percentage of total deaths due to cardiovascular causes than is typical.

**Ch. 6 (Conclusions on primary ozone NAAQS)**

This is generally a well-reasoned chapter in its attempt to ultimately motivate specific alternative recommendations for revisions to the ozone NAAQS. It is appropriately cautious when necessary. A few points merit comment.

1. Suggestion for more informative graphs.

In order to provide a more complete perspective to help judge the relative estimated impacts of alternative standards, I would like to see figures 6-1→6-6 (pp. 6-34 → 6-39) display the estimated decrease in effects of the progressively more stringent standards relative to recent

levels for 2002 and 2004, rather than only compared to just meeting the current standard (84/4). Exposures for many in the population are estimated to be reduced dramatically from recent exposure scenarios by just meeting the current standard.

2. Apparently puzzling (but probably correct) results of the risk-based approach.

Some of the results of the risk-based approach seem paradoxical and could benefit from some further explanation. For example, in Table 6-2 (p. 6-18) it is strange that twice as many individuals as are exposed to >80 ppb are estimated to experience a 10% fall in FEV1, when only 10% of population reacts this much at 60 ppb (vs. 25% at 80 ppb); therefore, most of these estimated falls in FEV1 must be due to exposure to <80 ppb. But, this is not what is shown by the "just meeting the standard" estimates, since these only result in decreasing the number estimated to have these FEV1 falls in half. Also, in Table 6-1 (p. 6-15), it seems strange that the % reduction in exposure due to more stringent standards is less in the worst case scenario (2002) than in the cleaner year (2004).

3. The form of the standard.

While it is painful to seriously consider substantially different forms of the standard, they might nevertheless be worthy of consideration. I have been impressed by the differences in population ozone exposures across cities (more so than for PM, for example) that have either always or sometimes been out of compliance with the ozone standard. For example, Denver has infrequent and relatively brief summertime episodes when concentrations skirt the ozone standard, sometimes resulting in Denver being out of compliance. In contrast, some cities experience sustained and repeated periods of elevated ozone concentrations resulting in chronic noncompliance. These contrasting patterns are not distinguished by the current form of the standard, yet likely have substantially different implications for the health of the populations in these cities. Why not at least consider some incremental form of the standard?

Another issue of relevance in considering a change in the form of the standard is the anomalies in estimated impacts on population exposures produced by meeting the current form of the standard (Table 4-9 [p. 4-32]). Specifically, there are marked disparities across cities, with dramatic effects on reducing estimated exposures in Houston and Los Angeles compared to those on other cities, simply due to the particular form of the standard.

On a small point, it is unclear to me how, in distinguishing it from the earlier form, the current form of the standard also "is not just whether the concentration is above a specified level" (p. 6-41).

4. Concentration-response and thresholds.

Evidence for thresholds based on controlled human experiments and epidemiological studies is qualitatively and quantitatively different, as reflected in the different ways of using this evidence in the risk-based approach.

At one extreme, it is argued (p.6-7 [line 18]) that it is difficult to find epidemiological studies that demonstrate associations at levels below the current standard, implying a threshold at levels that are only experienced in settings just in compliance with the current standard. On the contrary, I think it is easy to find studies that provide evidence that effects are estimated to occur in settings where the current standard would be met (most notably, see the Bell 2006 paper that is cited). Further, it appears that an attempt is being made to argue that because studies are done in areas that would likely not have been in compliance with the ozone standard (pp.6-9 → 6-11),

we can't know whether the findings of these studies would have been seen if the studies had been done when the standard was met. In fact, it is almost certain that most of the data that is responsible for estimating the effects seen is from days when ozone concentrations are far below the standard. When analyses are done on subsets of data that seem to have clearly met the standard, results are unchanged (again, see Bell 2006). Further, there are in fact studies that have been carried out when the current standard was not exceeded: for example, the Vedal et al (2003) mortality time series study from Vancouver, the Brauer et al studies of berry pickers (1996 and 1999 [J Environ Med]), and some respiratory hospitalizations studies (alluded to in Ch. 3-17 [lines 13-14]). And, likely there are years in some of the time series studies when the standard is met, even though, using the full study period, the standard would not be met. Therefore, there is ample evidence that associations in observational studies are present below the current standard. The real issue is one of how we interpret these associations, not whether they are present.

At the other extreme, it is stated (p.6-6 [line 31]) that the human experimental data provide evidence for a threshold at "near the *lower limit* [my italics] of ambient ozone concentrations in the US (CD, p. 8-44)." This is not correct. The statement needs to be dropped, or at least clarified if something else is intended.

5. The level of the standard and the level of certainty.

An argument is made that we are less certain of the findings of effects at lower ozone concentrations, and that this uncertainty should inform our choice of alternative standards. It is clearly true for the human experimental findings that we are less certain of effects at lower concentrations (e.g., 40 – 60 ppb vs. 80 ppb and above). This is not necessarily the case for the epidemiological findings, at least based on the epidemiological studies alone. That is, we are not less certain because the studies themselves are less certain. I find little difference in the certainty expressed about findings in studies done in higher vs. lower ozone concentration settings. Our level of certainty concerning epidemiological findings is based on information largely from outside of these studies, and therefore takes on a Bayesian flavor. Specifically, there are many reasons for being uncertain about epidemiological findings at low concentrations that are not directly related to the findings of the studies themselves. Two of the most important are: 1) there is justifiable concern that there is little relevance of central site ambient ozone concentrations to individual exposures, with exposure concentrations being much lower and poorly (or at least variably) correlated with monitored concentrations, a problem that is likely more acute in low concentration settings, and 2) there is questionable plausibility to the contention that the concentrations to which people are actually exposed in these lower concentrations studies can be responsible for significant (or any, for that matter) adverse effects, most notably death.

My recommendation for the level of the standard is in the range of 0.060 to 0.070 ppm. This recommendation is informed by clear demonstration, both experimentally and observationally, of effects below the level of the current standard. In 1997, we were aware of effects in experimental settings at concentrations of 0.080, concentrations already below the current level of 0.084. My lower limit is informed by my level of certainty in study findings, both experimental and observational, of effects at lower concentrations. Within my preferred range, I would select a standard closer to 0.070 than 0.060. The basis for this preference is twofold: 1) the impracticability of imposing a standard where the vast majority of the country would be out of compliance, and 2) the risk analyses demonstrating the relative impacts of

meeting the current standard and of lower proposed standards, with the majority of the benefits to public health achieved merely by meeting the current level of the standard. After all, standards that result in achievable benefits to public health are much preferred to those that might achieve these benefits only in an imaginary world.

6. Suggestions for uncertainties/research needs.

The overwhelming deficiency is the paucity of experimental human data on lung function decline at concentrations below 80 ppb. Since these data are very influential in both the exposure analysis and the risk analysis, increasing the number of subjects in such experiments, and assessing the reproducibility of such responses in individuals, should be a priority.

In regard to point 3 (p. 6-46) of the uncertainties and recommendations for future research, which touches on exposure, I would emphasize the need to have better exposure estimates for the population potentially at risk of dying from ozone exposure. Currently, a case can be made that exposures here are too low to plausibly be responsible for the dire effects estimated in the time series studies mortality studies. Point 6 expands on this point well. On point 2, I would like to see more specific recommendations for figuring out how and whether exposure measurement error is having an effect on the shape of the concentration-response relationship.

7. Minor points

page

6-13. What of exposure measurement error in influencing our level of certainty?

6-30. Reference to the Peters pilot study on very short term effects of ozone on MI gives it too much play. Remember that the more prominent very short term effects of PM in the pilot study were not seen in a larger and better designed study by these same investigators (HEI report 2004), although they did not specifically report estimates of very short term effects of ozone in this later study.

6-42. Figure references need correcting.

# Dr. James (Jim) Zidek

**Revised Comments of the Staff Report 2nd Draft**
**James V Zidek, August, 2006**

## CHARGE QUESTIONS

My detailed comments respond to a number of the issues raised below and only summaries are provided here.

### O3 air quality information and analyses (Chapter 2):

1. To what extent are the air quality characterizations and analyses clearly communicated, appropriately characterized, and relevant to the review of the primary and secondary O3 NAAQS?

**The results are generally clear and very well communicated. In particular, the Report provides a good characterization of the concentration field of the current criteria metrics.**

2. Does the information in Chapter 2 provide a sufficient air quality-related basis for the exposure, human health and environmental effects, health risk assessment, and environmental assessment presented in later chapters?

**I remained concerned about the level of level of uncertainty associated with the estimates of the PRB field. A sensitivity analysis of the estimated levels would be desirable. In a statistical context, a coefficient of variation of 10% would be considered a realistic target for accuracy and if it were assumed that the GEOS-CHEM model met this standard, it would mean varying that background by 10%, guided by estimates of that accuracy in the AQCD.**

### O3-related health effects (Chapter 3):

2. What are the views of the Panel on the appropriateness of staff's discussion and conclusions in Chapter 3 on key issues related to quantitative interpretation of animal toxicology and controlled-exposure human experimental studies and epidemiologic study results, including, for example, exposure error, the influence of alternative model specification, potential confounding or effect modification by co-pollutants, and lag structure?

My earlier concerns here have been adequately addressed.

### Exposure Analysis (Second Draft Chapter 4 of the O3 Staff Paper, draft Exposure Analysis technical support document, and OAQPS Staff Memorandum on Uncertainty Analysis):

1. To what extent are the assessment, interpretation, and presentation of the results of the exposure analysis as presented in Chapter 4 (and in the second draft Exposure Analysis technical support document) technically sound, appropriately balanced, and clearly communicated?

Generally the results are clearly communicated and well presented although the interpretation for setting standards of some of the APEX outputs is difficult. A number of specific concerns appear in the comments below.

2. Are the methods used to conduct the exposure analysis technically sound? Does the Panel have any comments on the methods used?
Based on my experience with a similar model, I believe APEX to be a sound methodology although too little assessment of its predictive accuracy has been carried out, especially given its central role in scenario analysis for standard setting.

3. To what extent are the uncertainties associated with the exposure analysis clearly and appropriately characterized in Chapter 4, the Exposure Analysis technical support document, and the uncertainty memorandum?
Overall, a very thorough well-balanced account of potential sensitivities and weaknesses of the models used has been given. In so far as these can be checked they seem correct and comprehensive. In some instances a quantification of uncertainty estimates is needed as indicated in my comments. More discussion and interpretation of the outputs is needed.

4. To what extent is the plan for the remaining uncertainty assessment technically sound? Are there other important uncertainties which are not covered? What are the views of the Panel on sensitivity analyses conducted to evaluate the influence of uncertainties in the exposure analysis?
This sensitivity analysis has been well done within the limitations of available data and the obvious sensitivity to AER has been discovered. However, the whole analysis of output uncertainty could have been carried out within the framework of a fractional factorial design to look for interaction effects between the different parameters. That might have been informative.

**Heath Risk Assessment (Second Draft Chapter 5 of the O3 Staff Paper and draft Health Risk Assessment technical support document):**

1. To what extent are the assessment, interpretation, and presentation of the results of the revised exposure analysis as presented in Chapter 5 (and in the second draft Risk Assessment technical support document) technically sound, appropriately balanced, and clearly communicated?
*These are generally sound and very well communicated.*

4. To what extent are the uncertainties associated with the health risk assessment clearly and appropriately characterized in both the second draft Chapter 5 and the second draft Health Risk Assessment technical support documents?
*The uncertainties are very well characterized. A central problem here and elsewhere derives from the problem of quantifying qualitative uncertainties such as model uncertainty.*

## GENERAL COMMENTS

**1** Overall the Second Draft Staff Report is well written.

**2** Some early general discussion of the ideas underlying the APEX approach should be included in the Staff document, given its central role in this exercise**.** After all, it is a big leap from the regulation of ambient levels to personal exposures. That discussion should in particular describe how later in Chapter 5 it is used in conjunction with an exposure response function and used in scenario analysis.

**3** Some concerns that need to be highlighted:

- The lack of assessment of APEX'S predictive accuracy. This seems surprising, given that this methodology along with its predecessor, pNEM, have been under development for well over a decade and a principal tool for such analysis.
- APEXs assumption that reducing ozone concentrations would not lead to behavioral changes. The Langstaff memorandum raises the same concern (**Langstaff page 56**).
- Uncertainty of accuracy of APEX predictions. Providing estimates of the uncertainty for things such as person-days of overexposure should be straightforward as APEX is a predictive probability distribution and "error bars" could have been generated and included in figures such as 5-11. **Note: page 6-16, line 31** promises such estimates in the Final draft.

**4.** The quadratic rollback procedure is not very well described in the AQCD, the Staff Reports or the supplementary material. Considering its vital role in the scenario analyses and exposure response modeling, I suggest expanding **lines 22-28, page 4-28** to include the formula, how it is used and the material on **page 5-10.** As now written, the reader learns for the first time in this document that it shrinks the actual ambient hourly concentration measurements into attainment.

**5** While the discussion of APEX uncertainties is remarkably thorough and comprehensive, they can only be partially quantified leaving concerns about the accuracy of the eventual predictions of such things as person – hours of exposure above a prescribed threshold. However, I see no practical solution with present knowledge. In future much more model assessment should be done to get a better handle on its accuracy.

## DETAILED COMMENTS

### Chapter 2:

**2-2 line 16:** The sentence beginning on this line should be the last sentence in its paragraph following the sentence defining $NO_y$.

**Appendix 2-A:** July (month 7) and August seem to have the lowest PRB and Sacramento has a large month to month variation. Some rationale should be included for what some reviewers of this document might find curious findings.

### CHAPTER 3

**3-36 Line 9:** Is 40ppb the "standardized unit" of change referred to in **line 10**? If so that same unit should be used again because the sentence is ambiguous as it stands.
**3-37 Line 3:** Should not this be "exposures to ambient concentrations"?
**Line 18:** This goes to a comment below and the 3rd concern above about APEX which assumes behavior would not change if ambient ozone concentrations were lowered.
**3-39 Line 8:** Some emphasis should be placed in future epidemiological studies in using a model like APEX to reduce errors in estimated population level exposures to the criteria pollutants, given the critical role APEX is playing in setting standards and given that it is freely available.
**Line 19:** I remain concerned about the observation reported in this line and ask that the rationale for it be included in the Report.

### CHAPTER 4

**4-1 Line 21:** The report should say something like "based on 2002 and 2004, for reasons given below in Section 4.3.4.1,…". However, two seems like too few years to really characterize the "year-to-year variation and more importantly bias, since 2004 had lower than typical O3 concentration levels.
**Line 24:** No doubt lowering ozone levels would lead people, especially susceptibles, to spend more time out of doors and thereby to sustain higher rather than lower exposures to ozone, a point not discussed in limitations of the APEX approach. (Runs condition on outdoor temperatures, not ozone levels, thereby missing this important interaction.)

**4-2 Line 28:** Something should be said here about "dose" (or a reference given). The text leaves the impression without explicitly saying so, that dose is important. Exposure (intake) is then estimated because it is an "important determinant" of dose. Why not estimate dose itself? If dose cannot be estimated, does the science at least say qualitative about the relationship between the surrogate (intake) and the thing it represents (dose) other than that the relationship is monotonic? For example is the relationship linear? That is would a 10% reduction in exposure lead to a 10% reduction in dose regardless of the level of that exposure?

**4-3 Line 8:** To avoid ambiguity, AQDC would be better than CD here and elsewhere in the report.

**4-4 Line 17**: Generally monitors are sited primarily to detect non-compliance and therefore, tend to overestimate ozone concentrations. Nevertheless the data they produce are used for other purposes as noted here in the report. However, it seems doubtful that they could be used to determine their own "representativeness" without the help of **Line 27.** Therefore these sentences should be linked in the text for clarity.
**Line 20:** For consistency the same units of measurement should be used throughout the report. Preferably, this should be ppb since numbers like 0.085 ppm for example, seem so small that seem unimportant.

**4-6** As noted in my earlier comments, APEX is a valuable tool for exposure assessment as well as predicting the benefits of a pollution abatement program. In particular, it can answer "what if" questions under roll back scenarios. However, APEX is only a model, an analogue of the real world therefore its outputs need to be assessed. I was pleased to see on page 4-13 Line 23 some recognition of this fact and a plan to do something about it in the future and at least a crude assessment on page **4-47.**

**4-11 Line 12:** The sentence added here in the revision seems to be a partial response to my comment on the first draft:
"randomness of model parameters needs some explanation in the Report. Their distributions could be interpreted by readers as Bayesian priors, reflecting (epistemic) model uncertainty, albeit with an empirical Bayes twist. In that interpretation, variability of the predictive output distribution would reflect that epistemic uncertainty as it propagates through the model, rather than only process (aleatory) uncertainty. It might be worth stating that these parameters are actually random effects so that variability in the predictive distribution represents aleatory uncertainty. Of course, that then pushes the issue of model uncertainty to a higher level on that staircase of infinite regress, that Mosteller and Tukey famously referred to. (See my comments below re Table 5.5.)"
Consider for example, the model in Equation (2-15) of "Exposure Analysis" technical attachment. The stdev's for it in Table 2-4 do reflect "uncertainties in the parameter estimates". In fact they are the standard errors of the estimates computed in a regression analysis of data from 32 clinical trials contrary to the assertion of that sentence. Moreover the issue of model uncertainty has been ignored even though it could well be the dominant source. In that model, for example, it seems quite plausible that a gender*age interaction term should be added. (In other words the coefficient b2 should be different from women than for men.)

**4-12 Line 7:** My earlier suggestion to interpret model outputs in terms of predictive probabilities and expected values has not been implemented. However that would make them easier to understand and make inter-city comparisons straightforward. Moreover, this approach would be more natural since APEX is a joint predictive probability distribution expressed as the product of various conditional distributions. Finally the use of probabilities and expectations in quantitative risk analysis is standard.

In any case the value of the city specific number of "person hours" in standards setting needs to

be clarified and their standard errors stated as they may not be negligible even if the number of APEX runs is large. (See comments on **page 53** below.)

4-14 Line 9: **No mention is made of the possible bias that might arise in the characterization of urban ambient ozone fields due to monitor placement for non-compliance detection. How were the monitors placed in the selected cities?**

**Line 21.** Using the multivariate spatial predictive distributions now available in modern space time process modeling, the variability of the random ozone field over the USA for a single or multiple years could well be captured and used in conjunction with APEX. (See Le and Zidek 2006. "Statistical analysis of environmental space-time processes." Springer.) This would overcome the difficulty associated with relying only on specified years.

That would not address the other kinds of uncertainties there regarding say changing demographic patterns and hence time activity patterns. This difficulty might be addressed partially by running APEX conditional on the demographic group as is done in pCNEM (referred to in my earlier comments). The predictive population estimates at that level could then be combined with different weights to determine the possible effects of demographic trends. Users of APEX could then predict exposures for population subgroups such as commuters or individuals with heart disease, by representing them probabilistically as a combination of demographic groups and combining the predictive estimates appropriately.

**4-15 Line 6:** The meaning of the sentence beginning on this line should be clarified. To many, "uncertainty" means "variability". Only things that do not vary are certain.

**4-18 Line 16:** As noted above, two is too few years to characterize variability due to year.

**4-21 Line 30:** Using LM=1 and LA =0 would seem more realistic than setting both equal to 0 as has been done. Anyway, rerunning with these alternatives would give some idea of the sensitivity of results to these choices.

**4-29 Line 4.** The caption could be misleading since "design value" commonly refers to a specified cut-off criterion. "The measured values of the current design metric for the modeled areas" would avoid the interpretation that different criteria are imposed on the different cities. By the way, I am not sure why the ratio in the last column was included. Would not the difference be preferable, if anything was really needed there? (By the way, notice how much simpler this table would be if ppb were used instead of ppm.)
**Line 11:** The graphs do not show the numbers but rather the percentages right in line with my suggestion above for page 4-12. These can be interpreted as the chances of the specified event occurring for a randomly selected individual from the specified sup-population.

**4-30 Line 5:** The figure captions are really complicated. How about inserting an example in the text to help the reader interpret them: "For example, Figure 4-2, shows a 42% chance that a randomly selected asthmatic child in Cleveland experienced repeated 8 hour exposures

exceeding 70ppb (Scenario 74/4) during 2002-2004 while undergoing moderate or greater levels of exertion. In contrast…."

**4-44 Line 20:** The sensitivity observed here poses a dilemma for developing an optimum National air quality standard since any one standard would result in varying exposures across the Country due to varying human behavior patterns. How should a National standard compromise between competing local optimum standards?

**4-45 Line 14:** the conclusion of "moderately sensitive" may need to be more clearly explained, given the –57% change we see in Table 4-15. The fundamental issue here is the reliability of the stochastic decay rate model. If it is deemed to be valid, choosing the $10^{th}$ and $90^{th}$ percentiles (the worst case scenarios) in place of the randomly distributed decay rate seems unduly severe. More natural: shift the center of the decay rate distribution to the $25^{th}$ and $75^{th}$ percentiles, thus preserving the uncertainty being expressed by that distribution while making it stochastically larger and smaller, respectively.

**4-47 Line 8:** Tailoring the model to fit the California situation assesses whether the model is ideally capable of accurately forecasting exposure. However, APEX as used to set National standards is a different model and it too should be assessed. Maybe it's AERs are higher than those in Sacramento and it does not underestimate exposure like its tailored counterpart.

**Line 18:** The shifty unit again, this time going from the $\mu g/m^3$ to ppm.

**ATTACHMENT: "Ozone population exposure analysis for selected urban areas".**

**10** Available predictive distribution methodologies (and software) would enable ambient levels to be randomly generated down to the tract level and thereby reflect the uncertainty in some urban areas about the ozone concentrations in the outside air being exchanged by the APEX model. Staff should consider this refinement. The result would be in line with the way APEX has been developed.

**12** The claim about the stability of A "over cohorts" seems surprising if this means "between" rather than "within cohorts". For example, one would have thought the autocorrelation for institutionalized seniors would be far higher than for young, working persons. **Page C-S12** states that auto-correlations within cohorts "may differ greatly" among its members. Some clarification is needed here. (I note that the model outputs are insensitive to "small changes" whatever that means exactly.)

**18** This page suggests, seemingly in contraction to **page 4-2,** that dose is being simulated in addition to inhalation. Some reconciliation of language would seem desirable.

**21** Clock hour "I" rather than "I" seems correct.

**22** Can some estimate be made of the effect of using the school children's D and A for all individuals? In particular is this path more liberal or conservative than that of drawing the sequence of daily time activity patterns at random?

**23** See the comment about **page 10** above re the imputation of local concentrations from those at the monitoring sites.

**40** There does not seem to be a **Section 3.8.2** so it was not clear what fractions were actually taken from **Appendix A** and used.

**46** Some discussion of the results in this section would have been helpful. How are the plots shown here to be interpreted and used, especially those about "person – days"? The same can be said about the next section, although there at least some discussion of the findings is given.

**53** It is not clear how to interpret the person-day estimates. More discussion is needed about how exactly these person-day measures of population impact are to be interpreted for setting standards.

No explicit statement is given about how the aggregates were computed. In the case of person-days, was this done by: first calculating the expected number of days of exposure per person based on a number of APEX runs; and then multiplying by sub-population size? If so the standard error of these estimates could be quite large even if a large number of APEX runs were made. Some statement of the reliability of these estimates needs to be given. A very relevant analysis of this reliability is given in the Langstaff memorandum.

**73** While this assessment is valuable and encouraging, nevertheless the suggestion that APEX underestimates true exposure is of concern. So is the use of weekly the aggregation of exposures. These could give quite a misleading impression of hourly exposures because of the ecologic effect.

**84** The technical appendices that follow on from this page seem very comprehensive and thorough. Where I was able to check the details, I found them correct. Moreover, relative to the complexity of their content, they seemed commendably clear.

**A-1** No information is given about the design of the survey used to select the 37 residences in the (smallest) RTP study. Nothing is said about the weights if any used in the estimates to adjust for unequal selection probabilities. Could the eventual estimates of population parameters be biased?

**E-6** I would have thought the regression model on this page should have been fitted after logarithm transformation and the resulting model on this page expressed at a product. That would have had the advantage of insuring amongst other things that the predicted AER is positive, whereas it could in principle be negative as presented.

**DRAFT MEMORANDUM: "Analysis of uncertainty in ozone population exposure modeling" by John Langstaff**

This memorandum is a generally clear and well-written account of some very technical issues. Some specific comments for the specified pages follow:

**2** **"**Variability" also leads to "uncertainty" so nowadays these two are commonly referred to as "aleatory uncertainty" and "epistemic uncertainty". The first represents the uncertainty that derives from the various random sampling processes built into the APEX simulator. The second, the more subtle of the two, represents lack of knowledge about the underlying stochastic models. In APEX this means uncertainty about the stochastic model parameters such as the one reflecting the variation in AERs. However, it should also reflect uncertainty about that model, say whether it is lognormal or any of the logstudents-t distributions with their varying degrees of freedom.

Why include model uncertainty? The answer: to buy insurance against model mis-specification that can lead to APEX outputs that are biased away from the actual exposures. Better to have a heavier tailed distribution that covers the exposures the model aspires to predict than an optimistically narrow one that misses the boat.

Incidentally, a proper assessment would determine the calibration of the APEX prediction intervals. Too much uncertainty would mean these say 95% intervals would be too broad relative to the observations and hence cover more than 95% of the observed exposures. That would point to the need for additional knowledge and hence reduction in epistemic uncertainty. **Pages 4 and 8** provide good discussion on this general issue about uncertainty.

**3** **2nd last paragraph:** Shouldn't this be "between input uncertainty" rather than "between model input uncertainty"?

**8** **Log-normal distribution.** Analysts usually log – transform data that are lognormally distributed, perform inference on that scale and then transform back. This leads immediately to confidence intervals for the GM and GSD since these are by definition, monotonic transformations or their transformed counterparts: GM = exp (AM) and GSD = exp (SD). This fact would lead to asymmetric confidence intervals for both. For the geometric mean for example, the 95% confidence limits would be $(GM/GSD^{2,} GM*GSD^{2})$.

Since the normal theory estimates, AM and SD, are stochastically independent, I am puzzled by Figures 17 and 18 that show them to be correlated. Is a different definition of GSD being used? (The mean and standard deviation of the lognormal distribution are related in a mathematically explicit way.) If so that definition should be stated for clarity. Or could this mean the lognormality assumption is inappropriate?

On this page and elsewhere bias in the ambient monitoring measurements in urban areas has been ignored.

**9** Indeed spatial variation in ozone concentrations can be considerable. For that reason, SHEDS has been applied with the pollution field interpolated down to the tract level. Obviously that would be impractical for APEX however.

**10** As noted above, the quadratic roll back process is not very well described anywhere in the documentation including the Staff paper.

**13** **First line:** Typo: "at a given monitor as uncorrelated".

The imputations of missing values have quantifiable errors that could be quantified and reflected in the uncertainty distribution for APEX outputs. However, the method proposed as an alternative seems satisfactory.

Incidentally, that method is the "cross-validation" approach rather than the "jack-knife" approach. The latter and its competitor, the "boot-strap" are used to correct bias in estimators and estimate their standard errors. That term should be changed on this and subsequent pages.

**14** Would not including all the monitors lead to an underestimation not overestimation of the prediction errors? As noted in the last line some of these will be zero.
**15** Why is the ratio rather than difference used as a measure of prediction accuracy. The former unlike that latter would have a very non-normal distribution.

The cross-validatory assessment should have been made by removing and not using the target site in constructing the spatial interpolator to get a better indicator of prediction error.

The basis for the assumption made at the end of the first paragraph on this page seems tenuous. Further analysis would eliminate the need for it.

A well-tested method for spatial prediction has been developed for space time fields and software is freely available online (http://enviro.stat.ubc.ca also includes a demo). The method overcomes the deficiency in geostatistical methods that ignore the stream of temporal data available for estimating the various parameters required for the interpolator. It also avoids the assumption of covariance stationarity made in kriging.

Incidentally, kriging, being a linear method, de facto assumes the underlying field is Gaussian. Its performance for ozone can therefore be improved by taking a square root transformation of the field and transforming back after completing the interpolation.

**21** APEX averages the working and home district ambient levels only in as much the microenvironments through which the sampled composite commuters go from home to work to home districts during weekdays. The first paragraph in the subsection beginning on this page gives a misleading impression that the two ozone fields are directly averaged. Furthermore the claim about ambient levels used for commuters who work outside the study area disagrees with what is claimed on **page 4-21** of the Staff Report. Reconciliation is needed here.

**22** The section beginning on this page does not make clear what uncertainties associated

with roll - back are being assessed. Presumably the reductions called by the quadratic formula are not uncertain. Is it the quadratic method? Is the uncertainty about how well this approach actual replicates what would happen under a real rollback?

**First sentence, 2nd paragraph of section.** Shouldn't "reflect" be changed to "attain"? In this same sentence, two sentences address how uncertainty is to be assessed by the proposed method. Is one redundant? Should they be merged? In any case, I could not understand why the first one is true. Why do the differences between them reflect the uncertainty that is being considered in this section?

**3rd paragraph of section.** If I understand its intended meaning correctly, the sentence beginning "A 3-year.." should be rewritten as: "A 3-year period with wide range ozone concentrations will typically have higher population exposures than one with a narrow range even though both periods have the same design value (or both are just in attainment)."

**34** Why are these values constrained to lie between 0.95 and 8.05? Why is that even mentioned at this point since the actual estimates are well within that range?

## CHAPTER 5

**5-18 Line 20 - :** This imaginative but somewhat ad hoc approach is not quite Bayesian because of the way in which synthetic data are created. A likelihood based alternative:

# Quantifying Response Probability Uncertainty

Jim Zidek

August 29, 2006

## 1 Credibility Bands

Suppose the number of responses $S$ among $N$ subjects at any given exposure level $x > 0$ has a binomial distribution with a logistic response probability

$$y(x;\alpha,\beta) = \frac{\alpha e^{\gamma}(1-e^{-\beta x})}{(1+e^{\gamma})(1+e^{-\beta x+\gamma})}$$

where $x > 0$, while $\alpha, \beta, \gamma > 0$ that is,

$$P[S=s] = \left( \begin{array}{c} N \\ s \end{array} \right) y(x;\theta)^s [1-y(x;\theta)]^{N-s},$$

for $s = 0, \ldots, N$ where $\theta = (\alpha, \beta, \gamma)$. Then given observed responses $s_1, \ldots, s_n$ at exposure levels $x_1, \ldots, x_n$ for subject totals $N_1, \ldots, N_n$, the likelihood for $\theta = (\alpha, \beta, \gamma)$ is

$$L(\theta) = \Pi_{i=1}^{n} \left( \begin{array}{c} N_i \\ s_i \end{array} \right) y(x_i;\theta)^{s_i} [1-y(x_i;\theta)]^{N_i - s_i}.$$

Suppose a prior distribution for $\theta$ that has a prior density $\pi(\theta) = \pi_1(\alpha)\pi_2(\beta)\pi_3(\gamma)$ where $\alpha \sim LN(\omega_1, \sigma_1)$ (the lognormal distribution), $\beta \sim LN(\omega_2, \sigma_2)$ and $\gamma \sim LN(\omega_3, \sigma_3)$. (For simplicity, the $\omega$s and $\sigma$s can be set equal to their maximum likelihood estimates to get an "empirical Bayes procedure.) Then the posterior (joint) density function of $\theta$ is

$$\pi(\theta|data) \propto \pi(\theta) \Pi_{i=1}^{n} \left( \begin{array}{c} N_i \\ s_i \end{array} \right) y(x_i;\theta)^{s_i} [1-y(x_i;\theta)]^{N_i - s_i}$$

where $data = \{s_1, \ldots, s_n\}$

We can sample items from the posterior using an MCMC approach that leads to a chain $\theta_1, \theta_2, \ldots$ which after a "burn-in" period, say a thousand iterations, becomes the required sample. The chain is Markovian so each successive element of the chain is conditional on its sequel. Thus given $\theta_i$, draw $\theta_{i+1}$ in steps by drawing $\alpha_{i+1}$, $\beta_{i+1}$, and $\gamma_{i+1}$ one at a time in order.

To draw them, we first need to choose a proposal distribution for each of $\theta$'s 3 coordinates. Thus for example we sample $\log \alpha_{i+1}$ from $N(log\ \alpha_i, \tau_1)$ where $\tau$ needs to be specified. (More below.) More precisely, generate $\log \alpha_* \sim N(log\ \alpha_i, \tau_1)$ and set

$$\log \alpha_{i+1} = \begin{cases} \log \alpha_* & \text{with prob } \delta \\ \log \alpha_i & \text{with prob } (1-\delta) \end{cases}$$

where

$$\begin{aligned} \delta &= min\{1, \frac{\pi(\alpha_*, \beta_i, \gamma_i|\text{data})q(\alpha_i|\alpha_*)}{\pi(\alpha_i, \beta_i, \gamma_i|\text{data})q(\alpha_*|\alpha_i)}\} \\ &= min\{1, \frac{\pi(\alpha_*, \beta_i, \gamma_i|\text{data})\alpha_*}{\pi(\alpha_i, \beta_i, \gamma_i|\text{data})\alpha_i}\}, \end{aligned}$$

where $q(\alpha_*|\alpha_i)$ represents the density of $\alpha_*$ conditional upon $\alpha_i$, that is

$$q(\alpha_*|\alpha_i) = (\sqrt{2\pi}\tau\alpha_*)^{-1} e^{\frac{(\log\alpha_* - \log\alpha_i)^2}{2\tau^2}}.$$

Ideally $\tau_1$ should be chosen to make the acceptance rate just under 50%. Draw $\beta_{i+1}$ in a similar fashion after replacing $\alpha_i$ by $\alpha_{i+1}$. Then $\gamma_{i+1}$.

As the sequence above is being generated, the sequence $\{y(x;\theta_i\}$ may be also be generated for any fixed exposure level say $x = 0.050$ after the burn-in period. Making a histogram or equivalent probability density estimate from the latter sequence will reveal the shape of that density. At the same time finding its $5^{th}$ and $95^{th}$ percentiles will yield an estimated 95% credibility interval for the response probability.

Repeating the procedure above for say 20 different exposure levels, $x$ and smoothly joining the upper limits and then the lower limits of the credibility bands will lead to a error bands around the estimated response curve.

## 2 Model Uncertainty

Usually model (epistemic) uncertainty will be substantial and the credibility band computed above, based on a single assumed and fixed model, will be too narrow. In fact, the true model is completely unknown save perhaps for the fact that it must be monotone increasing. Moreover, the very limited amount of data available at the left hand end of the range of exposure levels is too small to serve as a useful guide to a parametric form. Thus it would be preferable to redo the above analysis for a range of plausible models, including the logistic and linear ones. The latter would be given by

$$y = \alpha + \beta x$$

where $y$ denotes the response probability, $\alpha \sim N(\omega_4, \tau_4)$ while $\beta \sim N(\omega_5, \tau_5)$ and, again for simplicity, the requisite parameters again use the values already obtained for the current document using by maximum likelihood.

An additional model of particular interest would be the *broken stick* that allows a threshold $\mu_1$ to be fitted at the lower end of the range:

$$y = \begin{cases} \beta_1(x-\mu_1) & \text{for } \mu_1 \le x \\ \beta_1(\mu_2-\mu_1)+\beta_2(x-\mu_2) & \text{for } \mu_2 \le x < \mu_3 \\ \beta_1(\mu_2-\mu_1)+\beta_2(\mu_3-\mu_2)+\beta_3(x-\mu_3) & \text{for } \mu_3 \le x \end{cases} .$$

Again lognormal distributions should be used for the $\beta$s and normals for the $\mu$s although with 6 parameters this model will be the most difficult to fit with maximum likelihood. In more elaborate analysis it would be penalized for its lack of parsimony but that refinement might be left for the future in view of time constraints on the process.

How can the uncertainty about these (and others) models be combined? One simple approach would be to weight them according to biological plausibility (which makes the linear model quite unrealistic at the right hand end at least) say

| Model | Weight |
|---|---|
| logistic | 0.50 |
| logistic | 0.30 |
| brokenstick | 0.20 |

Running the MCMC analyses for the posterior distributions associated with each of the models will yield sequences say of 10,000 iterations after burn-in, $\{y_j(x;\theta_1^j), y_j(x;\theta_2^j,\ldots)\}$ for model $j$ = logistic, linear, brokenstick. The 10,000 values for the logistic model can then be combined with the last 6,000 and 4,000 values for each of the linear and broken stick models, respectively, to get a single combined set of values. Its arithmetic mean gives an estimate of the combined response probability while its $5^{th}$ and $95^{th}$ percentiles would yield a 90% credibility band for that unknown response.

At the lower end of the range of exposure levels this band would be wider than the one for the logistic model alone, reflecting *epistemic uncertainty*. Its large size would reflect both the greater *aleatory* and epistemic uncertainty there, the former due to the lack of information that can be gleaned from the data available at low exposure levels.

In personal communication, Dr Lianne Shephard noted that the large number of parameters in the broken stick as formulated would make it difficult to fit with the relatively small amount of data, emphasizing anew the uncertainty at the left hand end of the curve. She also made suggestions that improved the exposition in this technical section of comments.

**5-21 Line 28.** A very good description of *variability* or what is commonly referred to as aleatory uncertainty.
**5-22 Line 20:** Word missing here.

## CHAPTER 6

**6-13 Line 6:** A potentially important bias could arise from the changes in the time-activity patterns of members of the susceptible groups, such as those used to characterize the exposure – response function and as noted above, this source should be included in the list of uncertainties.

**6-34 Figure 6-1:** In this and the falling figures we see a surprising amount of variability between cities for 74/5 compared with 74/4 and that is less than 74/3. That seems curious. In any case, it would argue for 74/4 on the basis of robust of the predicted benefits across cities.

**6-47 Line 1:** Environmental epidemiologist could well use APEX – type models in their assessments of association of health risks and pollution. Such models in conjunction with interpolation of ambient fields to local areas could help overcome the measurement error effect in the use of ambient concentration fields as surrogates for exposure. It would also provide an estimate of the exposure response as against the concentration response. Moreover, based on my research work with an APEX relative (pCNEM) non-significant associations can be turned into significant associations through the de-attentuation of the relative risk.

# Dr. Barbara Zielinska

## Comments of the 2nd Draft Ozone Staff Paper
## Chapter 2

**Barbara Zielinska**
**DRI**

In overall, this second draft of the ozone Staff Paper represents significant improvement over the 1st version. Below are my responses to the charge questions to the CASAC $O_3$ panel from Dr. Martin's memo (July 17, 2006), regarding Chapter 2.

1. *To what extend are the air quality characterizations and analyses clearly communicated, appropriately characterized and relevant to the review of the primary and secondary $O_3$ NAAQS?*

In my opinion, Chapter 2 is well written and presents a concise summary of the information contained in Chapter 2 and 3 of the final $O_3$ AQCD. The ambient ozone levels, $O_3$ precursors, their sources and emissions, temporal and spatial variability, long-term trends, and characterization of ozone episodes are adequately summarized in this chapter. This information is relevant to the review of primary and secondary $O_3$ standards.

2. *Does the information in Chapter 2 provide a sufficient air quality-related basis for the exposure, human health and environmental effects, health risk assessment, and environmental assessment presented in later chapters?*

Section 2.2.6 covers briefly the relationship of ozone to other photochemical oxidants. However, there is no information in this chapter regarding the role of ozone and other photochemical oxidants in the atmospheric transformation processes that may results in the formation of more toxic products (both in outdoor and indoor environments), either in gas or particle phases. This topic, covered in $O_3$ AQCD, seems to be relevant to human health and environmental effects.

The policy relevant background (PRB) section (2.7) continues to have problems. Although this section very briefly cites the results of the comparison of different models and measurements, it does not adequately address the uncertainties of the global GEOS-CHEM model and how these uncertainties are reflected in the health risk analysis. The authors say (page 2-48, line 1-4) that the GEOS-CHEM model overestimates PRB ozone values for afternoon surface air over southeast in summer and it is accurate within 5 ppbv in other regions and seasons. If the PRB concentrations are predicted to be generally in 15 to 35 ppb range, 10 ppb overestimation is substantial. Besides, it is not clear how this information was derived. Since ozone health effects are observed down to concentrations of the order of 0.04 – 0.05 ppm, it is important to know how the PRB compares to the considered primary ozone standard and what uncertainties there are in the risk attributed to controllable sources.

Minor comments:

1. Page 2-2, lines 16 -21. The order of sentences should be changed. The sentence “NOx is considered a good surrogate for NOy” should be the last sentence, after NOy is defined.
2. Table 2-1. “Total without fire” for 2003 and 2004 should be moved up one line.
3. Page 2-5, line 30. Formaldehyde should be spelled out (H2CO may not be clear to everybody)
4. Page 2-15, lines 27-29. Transport is not evident from Figure 2-6
5. Page 2-48, line 8: The text comments very briefly that the PRB O3 concentrations are the highest during spring and decline into summer. However, as evident from Appendix 2A, diurnal ozone PRB profiles are very different for some city, for example for Houston or Sacramento. Any explanation?

# NOTICE

This report has been written as part of the activities of the U.S. Environmental Protection Agency's (EPA) Clean Air Scientific Advisory Committee (CASAC), a Federal advisory committee administratively located under the EPA Science Advisory Board (SAB) Staff Office that is chartered to provide extramural scientific information and advice to the Administrator and other officials of the EPA. The CASAC is structured to provide balanced, expert assessment of scientific matters related to issue and problems facing the Agency. This report has not been reviewed for approval by the Agency and, hence, the contents of this report do not necessarily represent the views and policies of the EPA, nor of other agencies in the Executive Branch of the Federal government, nor does mention of trade names or commercial products constitute a recommendation for use. CASAC reports are posted on the SAB Web site at: http://www.epa.gov/sab.