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

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

June 21, 2011

EPA-CASAC-11-008

The Honorable Lisa P. Jackson
Administrator
U.S. Environmental Protection Agency
1200 Pennsylvania Avenue, N.W.
Washington, D.C. 20460

Subject: Consultation on EPA's *Ozone National Ambient Air Quality Standards: Scope and Methods Plan for Health Risk and Exposure Assessment* (April 2011) and *Ozone National Ambient Air Quality Standards: Scope and Methods Plan for Welfare Risk and Exposure Assessment* (April 2011).

Dear Administrator Jackson:

The Clean Air Scientific Advisory Committee (CASAC) Ozone Review Panel (Enclosure A) met on May 20, 2011 to conduct a consultation on EPA's *Ozone National Ambient Air Quality Standards: Scope and Methods Plan for Health Risk and Exposure Assessment* (April 2011) and *Ozone National Ambient Air Quality Standards: Scope and Methods Plan for Welfare Risk and Exposure Assessment* (April 2011). The CASAC uses consultation as a mechanism for individual technical experts to provide comments to guide the Agency on issues early in the development of a document, before the first draft is ready for peer review. While we are not providing consensus comments on these documents, individual written comments on the two Plans are provided in Enclosure B.

A consultation is conducted under the normal requirements of the Federal Advisory Committee Act, which include advance notice of the public meeting in the Federal Register. As this is a consultation, we do not expect a formal response from the Agency. We thank the Agency for the opportunity to provide advice early in the NAAQS review process, and look forward to the review of EPA's First Draft Risk and Exposure Assessments on Ozone.

Sincerely,

***/signed/***

Dr. Jonathan M. Samet
Chair, Clean Air Scientific Advisory Committee

Enclosures

## NOTICE

This report has been written as part of the activities of the EPA's Clean Air Scientific Advisory Committee (CASAC), a federal advisory committee independently chartered to provide extramural scientific information and advice to the Administrator and other officials of the EPA. CASAC provides balanced, expert assessment of scientific matters related to issues 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 within the Executive Branch of the federal government. In addition, any mention of trade names or commercial products does not constitute a recommendation for use. CASAC reports are posted on the EPA Web site at: http://www.epa.gov/casac.

**Enclosure A**

**U.S. Environmental Protection Agency**
**Clean Air Scientific Advisory Committee**
**Ozone Review Panel**

**CHAIR**
**Dr. Jonathan M. Samet**, Professor and Flora L. Thornton Chair, Department of Preventive Medicine, University of Southern California, Los Angeles, CA

**CASAC MEMBERS**
**Mr. George A. Allen**, Senior Scientist, Northeast States for Coordinated Air Use Management (NESCAUM), Boston, MA

**Dr. Joseph D. Brain**, Cecil K. and Philip Drinker Professor of Environmental Physiology, Department of Environmental Health, Harvard School of Public Health, Harvard University, Boston, MA

**Dr. H. Christopher Frey**, Professor, Department of Civil, Construction and Environmental Engineering, College of Engineering, North Carolina State University, Raleigh, NC

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

**Dr. Helen Suh**, Program Area Director, Environmental Health, Public Health Division, National Opinion Research Corporation (NORC) at the University of Chicago.

**Dr. Kathleen Weathers**, Senior Scientist, Cary Institute of Ecosystem Studies, Plant Science Building, Millbrook, NY

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

**Dr John Bailar**, Scholar in Residence, The National Academies, Washington, D.C.

**Dr. Michelle Bell**, Associate Professor, School of Forestry and Environmental Studies, Yale University, New Haven, CT

**Dr. David Chock**, Independent Consultant, Bloomfield Hills, MI

**Dr. William Michael Foster**, Professor, Pulmonary and Critical Care Medicine, Duke University Medical Center, Durham, NC

**Dr. Judith Graham**, Independent Consultant, Pittsboro, NC

**Dr. David Grantz**, Director, Botany and Plant Sciences and Air Pollution Research Center, Riverside Campus and Kearney Agricultural Center, University of California, Parlier, CA

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

**Dr. Daniel Jacob**, Professor, Atmospheric Sciences, School of Engineering and Applied Sciences, Harvard University, Cambridge, MA

**Dr. Steven Kleeberger**, Professor and Lab Chief, Laboratory of Respiratory Biology, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, NC

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

**Dr. Howard Neufeld**, Professor, Dept. of Biology, Appalachian State University, Boone, NC

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

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

**Dr. Peter Woodbury**, Senior Research Associate, Department of Crop and Soil Sciences, College of Agriculture and Life Sciences, Cornell University, Ithaca, NY

**SCIENCE ADVISORY BOARD STAFF**

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

**Enclosure B: Individual Comments on Scope and Methods Plans**

# Table of Contents

**Comments on *Ozone National Ambient Air Quality Standards: Scope and Methods Plan for Health Risk and Exposure Assessment (April 2011)***

**Mr. George Allen**

Section 2.2, Pg 2-1, 2-5, and elsewhere throughout this document:
For the REA aspects of this analysis it is noted that the most recent three years of ambient ozone data (2008-2010) will be used, while GEOS-Chem modeling will be for 2006-2008. There are two major issues with these date ranges. First, it just makes sense to use the same time period for both - e.g., use 2006-2008 for the REA. Or, if that's not sufficient, run the REA for both periods.

Running the REA for 2008-2010 has its own complications. My concern here is that any one 3-year period may have unusual meteorological conditions in one or more of the years. In this case, 2009 was very notable for a cool and wet summer in much of the eastern US, resulting in distinctly lower seasonal ozone conditions. Region I posts trend analysis on their ozone web page at: http://www.epa.gov/region1/airquality/standard.html. One example for all of New England is Figure 1, from: http://www.epa.gov/region1/airquality/images/NE8HR.gif:

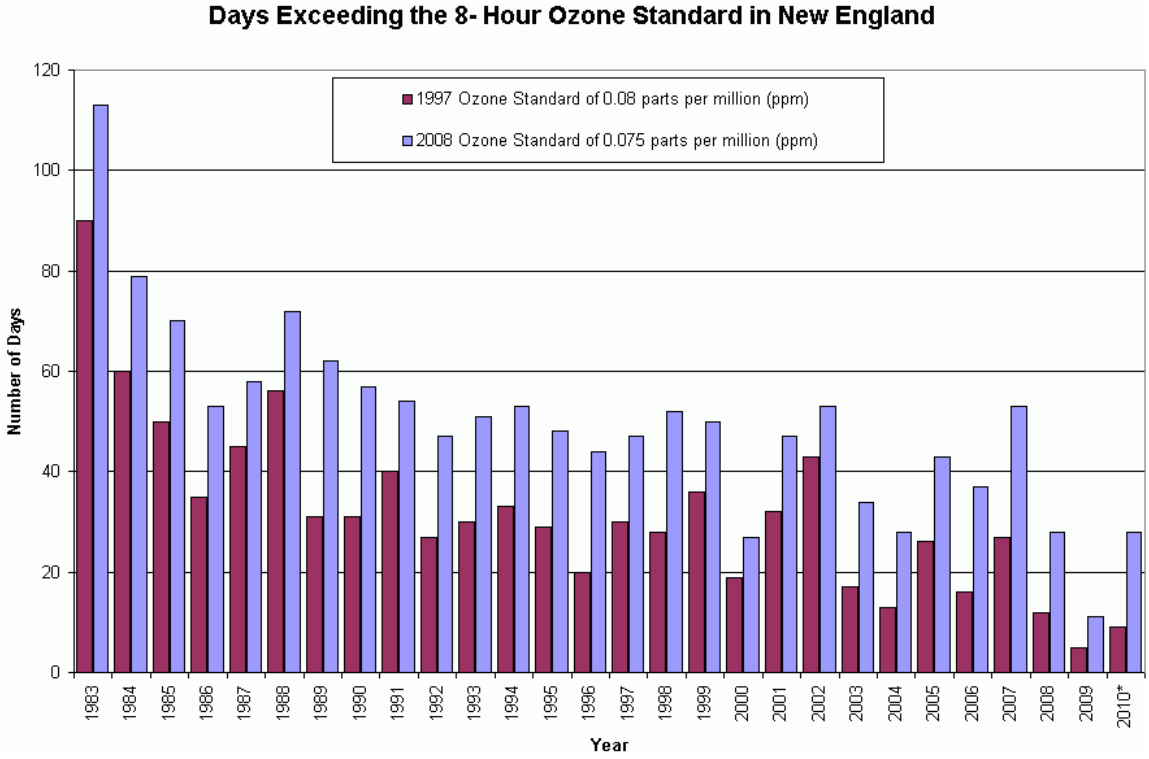


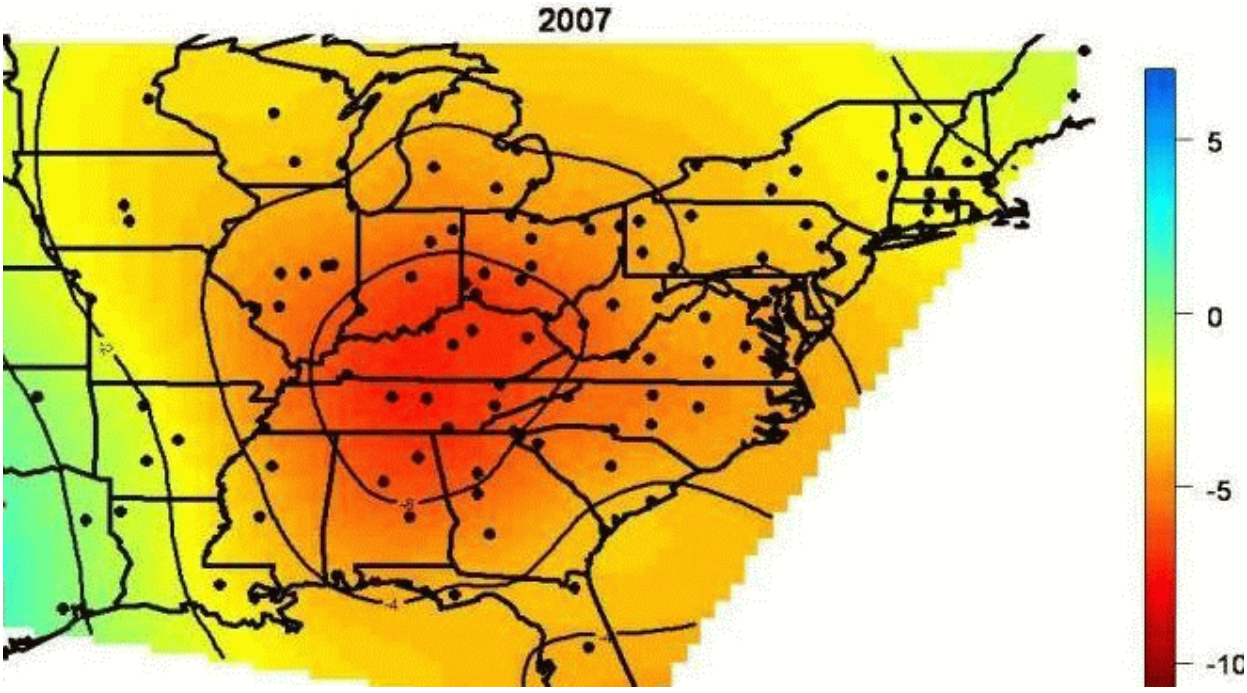

Figure 1.

Note that 2009 is the lowest since 1983, and that at the75ppb threshold, 2009 had fewer than half the number of days than the next lowest year. This chart also shows 2003-2004 as "clean years" and 2007 as a "not clean" year. Both 2008 and 2010 were normal to cleaner than normal years. In cleaner areas like VT (http://www.epa.gov/region1/airquality/images/VT8HR.gif) both 2009 and 2010 (preliminary data) had no days over the 75ppb standard; the only other year with no days was 2006.

Thus, a three-year metric for 2008-2010 may have an unusual low bias in the eastern US due primarily to meteorological conditions. A more through analysis than that presented here should be done of course.

When EPA does ozone trend analysis, a meteorological correction is applied to minimize the effect of unusual years on the trend (Camalier et al., 2007). There is substantial year to year variation in various ozone metrics in many areas of the US, but it is a reasonable assumption that sources (ozone precursors) do not change much from year to year, and if they do it's usually a monotonic downward trend as various control strategies are implemented. Figure 2 shows the adjustment to the ozone trend for 2007, a year with higher than usual ozone levels over much of the eastern US. A similar figure but with the opposite correction is at:
http://www.epa.gov/airtrends/weather/adj_2003_map.jpeg

Figure 2 (source: http://www.epa.gov/airtrends/weather.html).




These two years (2003 and 2007) have a correction range of more than 10 ppb. Unusual years like these can influence a 3-year ozone metric.

This document does not mention if the 2008-2010 ozone data will be corrected for meteorology or not. While one can make the case either way, it is likely that metrics for these three years will be lower than expected in parts of the country due to meteorological conditions. EPA should include this discussion in their Plan, and state why or why not these kinds of corrections should be made to the ambient ozone data that will be used in the analysis.

Section 2.3.
The use of quadratic rollback is only briefly explained. A more complete explanation and the impacts of the different rollback approaches is needed. The issues raised in the CASAC ozone panel's comments on the "First Draft Ozone Staff Paper, Risk Assessment, and Exposure Assessment Documents" in February 2006 are still relevant.

Section 2.3.1, pg. 2-4, line 5 mentions discussions in section 2.3.3, but there is no such section in this document. Presumably it should read 'section 2.3.2'.

Section 5 discusses issues related to spatial heterogeneity of ozone across urban study areas (Pg 5-9, lines 17 and 19; Pg 5-10, line 6) . With the existing urban data available, many cities do not have sufficient monitoring sites to properly address this issue directly. Boston, one of the candidate cities for this analysis, only has one site in the urban core (Roxbury). There are two other sites that may be relevant (Long Island in Boston Harbor, and Lynn ~12 miles NE of downtown), but neither site can be considered to reflect ozone within ~3-5 miles of the core area of Boston. These other 2 sites are very similar for 8-hour ozone data, and both somewhat different from Roxbury. Notably, there is a ~9 ppb intercept in the regression of Lynn vs. Roxbury for 2010 data. This likely reflects the impact of NO titration at the Roxbury site that is minimal at the Lynn site. Figure 3 shows a scatter plot of Lynn vs. Roxbury 8-hour ozone.

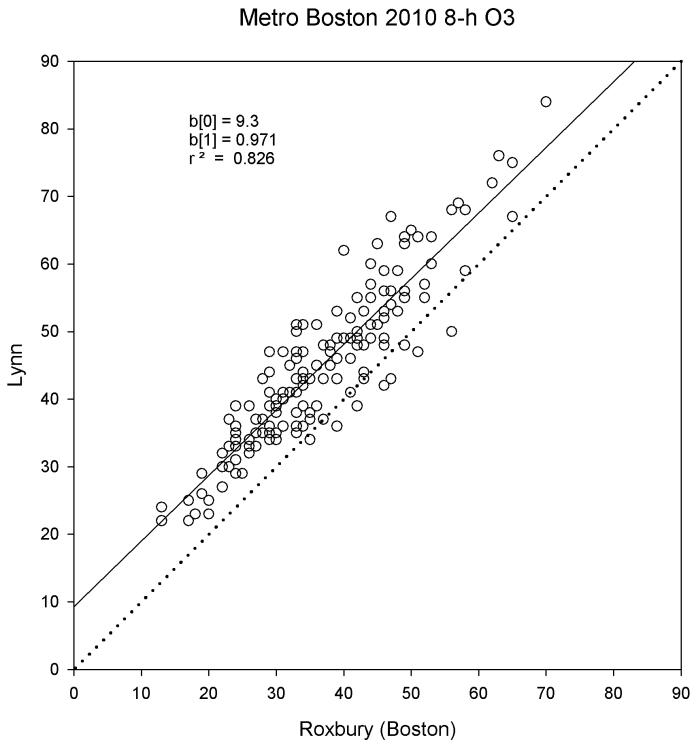



Figure 3

The effects of spatial heterogeniety within a metropolitan area are usually considered to impact the slope of the relationship between two sites; what implications does an offset such as this have in the planned analysis?

Note that as expected, ozone at the suburban Lynn site is generally higher than Roxbury, and never significantly lower. This document discusses the need for reasonable spatial correlations across an urban study area, but does not discuss cases like this where correlation and slope can be good but with a substantial offset across sites. The question would be: does this Boston scenario influences the kinds of analysis approaches discussed in Chapter 5.

References.

Camalier, Louise; William Cox, and Pat Dolwick (2007). The effects of meteorology on ozone in urban areas and their use in assessing ozone trends. Atmospheric Environment Volume 41, Issue 33, October 2007, Pages 7127-7137 doi:10.1016/j.atmosenv.2007.04.061

**Professor Ed Avol**

P1-7, Sec 1.2 Goals for Framing the Assessments in the Current Review, line12: "…impaired lung function…" – This raises an important issue, namely the definition of "impaired" lung function. Is it clinically significant? More than a 20% change? An "adverse" health effect? Exposures to very low levels of ozone can result in measurable changes in lung function in a large segment of the population, depending on ventilation rate, duration of exposure, etc – for the purposes of this plan, where is the threshold of "impaired" being set?

P1-8, line16, "…2008-2010.." – how was this time frame selected and why? Were these three years "representative" – abnormally wet or dry, hot or cold, El Nino or not? Or were they just the most recent years for which data was available?

P1-8, line 18, "…April to October…" – The assumption is that the photochemical ozone season is, was, and will be April to October. Given the concern over climate change and the wealth of meteorological information, is there any evidence that climate change has affected the "ozone season" timing?

P1-9, Section 1.3.2 Exposure Assessment, lines 5-8,"…the 12 urban cities modeled…" – no rural areas are apparently being considered in this analysis, but the exposure and time course of ambient concentrations can be very different in the urban compared to rural environment. In the urban case, atmospheric NOx titration from combustion products encourages a sharp O3 peak in daytime ozone (for example, in Los Angeles). In the rural case, without the local NO to quench the process, a broader sustained ambient level is often observed (for example, Lake Arrowhead). This different pattern of exposure will lead to different exposures for the population at different times of day for different cumulative periods of the day…how will this be taken into account if only urban areas are considered in the analysis?

P2-1, Section 2.2.1 Recent Air Quality, lines 29-30, "…only data collected by FRMs or FEMS will be used…" – the QA aspect of this is understood and appreciated, but this seems to contradict the earlier stated desire for personal exposure data and personal sampling information (p1-10, lines 12-14 ("because data on personal hourly ozone exposures are not available, estimates of personal exposures…")..

P2-2, lines 8-11 (use of composite monitors for data) – a composite monitor approach generally provides a better estimate of population exposure, especially for those freely moving through multiple microenvironments, so why wouldn't this be a preferred approach?

P3-1, Approach for Population Analysis, line 9, "…school-age children, ages 5-18…" - NIH considers subjects to be children up to age 21, and lung growth and development continues (for males) into their early 20s, so a slightly broader range of "children" may well be warranted. One could ask why truncate consideration of young children at 5 years, and not go younger. Infants and young children obviously breathe ambient air, spend time outdoors, and have high ventilation rates for their physical size, so may well be at higher risk. How can they be included?

One can make a similar argument for "elderly". Why make the cut at 70 years? Why not consider those over 65 years, which is more commonly used in the health care data?

P3-5, Figure 3-1, Item 7 (Calculate population exposure statistics) – two of the previously identified subpopulations of concern (elderly (>70 yrs) and outdoor workers) not listed here.

P3-7, Section 3.3, Populations Modeled, line 12 – school age children, aged 5-18…see previous comment above.

P3-8, line 7, “,,,1 occurrence for 10 people is counted the same as 10 occurrences for 1 person.” – timing and seasonality are potential issues here. Getting exposed one time at the start of the smog season (when few previous exposures have likely occurred and sensitivity is high) may exact quite a different response than getting exposed later in the season (for that portion of the population that develops an attenuated response upon repeated ambient exposures). Additionally, because this development of toleration or attenuation does occur for some individuals, 1occurence for 10 people does not universally equal 10 occurrences for 1 person.

P3-8, Section 3.5, Selection of Urban Areas and Time Periods – Because of NO titration in urban areas due to combustion exhaust, ozone is “peakier” in urban areas than the broader gradual rise and fall one often observes in non-urbanized locations. This leads to different patterns of temporal exposure, different cumulative exposures, and different temporal periods of importance for physical activity considerations (which becomes important for dose)…is there to be no consideration of non-urbanized locations?

P3-8, line23 – in view of climate change concerns, is there any evidence that the photochemical smog season has shifted temporally from the April/October time period?

P5-9, Assessment of spatial heterogeneity of ozone across urban study areas, lines 19-30 – the intra-urban ozone spatial pollution surface is likely to be more variable than previously appreciated, due to local NO titration near roads. This implies that variable exposures will occur, depending on spatio-temporal patterns of activity. Air monitors have historically been sited to avoid this local-near-road inhomogeneity, so data may not exist in the monitoring network to more accurately assign exposures -- so what can be done to improve upon this misclassification?

**Dr. John Bailar**

OZONE NATIONAL AIR QUALITY STANDARDS: SCOPE AND METHODS PLAN FOR **HEALTH** RISK AND EXPOSURE ASSESSMENT

Throughout, this document frequently says, "we may", "we are considering" and the like, but rarely says *how* EPA will decide whether to go ahead or not. This seems to me to be a serious gap.

Page 1-7. The list of goals is cast in terms of all-or-none. This is appropriate for mortality and some other endpoints, but not for quantitative matters such as pulmonary function. Maybe you have no way to deal with degree of impairment (including exacerbation of lung damage not due to ozone) but the matter should at least be addressed here and elsewhere in the document. (It appears that lines 5-14 on the next page do not cover this.)

EPA is well aware that short-term lab studies cannot inform us about long-term responses, but this should be acknowledged on pages 1-9/10

I simply do not understand the bullet that begins at the bottom of page 1-10. The third bullet is also unclear. In general, these four bullets may be correct, but they need substantial re-writing.

The statement in lines 5-7 on page 1-1. Not specific? Volunteers are generally drawn form local residents, and they may therefore carry the effects of any prior exposures with them. I would be pretty cautious about assuming that a person with past exposures to substantial levels of ozone or other pollutants would respond just like a person without that exposure, and I would not want to guess which direction such an effect, if any, might go.

Page 2-1, line 5. How is (1) different from (2) and (3)? Aren't they to be evidence-based to the extent possible?

Top page 2-3. Spatial and temporal heterogeneity within an area raises issues quite different from such heterogeneity across areas. This could use some discussion.

Page 2-5, lines 9-18. How were these grid sizes chosen?

Page 3-, lines 8-13. Please say more about how the areas will be selected, or insert a cross-reference to Section 5.3.2.

Sections 3-2 to 3-5 say a lot about what APEX *can* do, but Section 3-6 does not say enough about what will in fact be done with it.

Table 3-2 could usefully include a column that says how the study participants were selected. I am a little concerned about possible bias. The first bullet on page 3-16 does not really address the problem of non-representativeness *within* specific categories or demographic descriptors, but survey statisticians quite broadly recognize that anything short of a random sample with a high response rate may be seriously biased.

Averting behavior (page 3-17) is, in my view, less a scientific issue and more a philosophical/political issue. There should be at least brief mention of the tension centered on questions of how much sensitive (or other) persons should have to change their behavior to mitigate a hazard. (Similar questions have arisen in other situations, e.g. the broad agreement that it is inappropriate, and even illegal, to refuse employment to potentially sensitive persons such as women of childbearing age.) See also the top of page 5-16. One issue is how to accommodate the costs (broadly defined) of present averting behavior vs. the benefits (and costs) of future adoptions of averting behavior.

Several activities will be undertaken "if sufficient resources are available". But what is the priority ranking of these, if you can do some but not others? This seems to require some thought now, rather than later.

Page 3-19, line 7 refers to the joint distributions of parameters, but there is also reason to be concerned about the joint distribution of effects – i.e. synergy. This might be mentioned here.

I enthusiastically support the last paragraph on page 3-19.

Section 5.3.2 is an important aid in understanding this proposal, but I remain concerned about the use of 12 urban areas to represent the entire (urban?) population of the US. These 12 should be examined separately and compared. Only if they are in pretty good agreement should they be combined into a national figure, but I do not see a description of this critical comparison, including criteria for judging whether they are sufficiently similar to provide a reliable combined estimate. If they are not sufficiently similar, reasons for differences should be sought and, to the extent possible, the effects of differences should be estimated. This could be in the form of a sensitivity analysis, or perhaps by allocating each US city to a category that closely matches one or more of the 12 areas.

The totals on page 3-15 are a bit silly given the vast differences in quality and depth of information, and the fact that bias will surely be a bigger threat to conclusions than variance.

Equation 4-1. Using interval midpoints is appropriate only when the midpoint fairly represents the whole interval. In regions of a curve that are flat or changing evenly this may be OK, but it can lead to errors in categories near a point of inflection. This can be remedied by using sufficiently fine intervals, or by using a point a bit toward one side or the other within a broader interval (toward the higher side if the curve is bending downward, toward the lower side if it is bending up). A quadratic fit might also work.

Page 4-5, line 18. Children under 8 may present special problems. Do you recommend research to fill this gap the next time around?

In the equations on pages 4-4/4-5, are all of the betas constants to be estimated from the data?

Page 5-1, last line. Is the inequality reversed?

Page 5-2. I do not understand where the footnote is pointing because it is already in section 5-1.

Page 5-3, top. It is not clear how you will select the "modeling elements" to be taken as core vs. those for sensitivity analysis. If you will make the selection after you look at results, say so, with some

comment on how you will choose; otherwise give the reader some guidance about your choices prior to analysis. The paragraph on page 5-4 helps, but does not go far enough given the criticality of this phase of the analysis.

Footnote 5-1 belongs in the main text.

Page 5-8, figure. My copy of the draft document does not include color, but I infer that "Incidence and Prevalence Rates" is an input to "Adverse Health Effects". This seems backwards, unless "Incidence and Prevalence" refers to background rather that the effects of ozone. Please clarify. Also, I would give this figure a figure number.

Page 5-9, line 4. Can you possibly put these critical details in the next version of this document?

The paragraph at the bottom of page 5-11 is quite hard to follow, but it is worth noting that the evidence for a broad effect (e.g. total mortality) must be rated as being at least as strong as the evidence for any component effect (e.g. respiratory mortality) unless you admit the possibility of countervailing benefits in some other component. In short, if you decide that there is an effect on respiratory mortality, you will almost have to conclude that there is an effect on total mortality, whether that can be established independently or not.

Page 5-13, line 25. A finding that two or more effects in a multi-pollutant model are non-significant (not "insignificant", please) should be followed by computations that omit one, two, … pollutants to see whether they remain non-significant. When two pollutants are highly correlated in occurrence, the data may just be saying that one (or possibly both) is having an effect, but the data cannot tell which one it is. Please do not pass over a pollutant just because it is non-significant in a multi-pollutant model; there may be an important effect that is masked by the correlations.

Pages 5-22/23. I am not a fan of the WHO Tier classification because it gives zero guidance to the analyst. The words in the WHO descriptors of tiers are subject to a vast range of interpretations and distortions, and there are no sharp divisions among the tiers anyway. It is useful only for describing what was done after the fact. This bit should be deleted or replaced with a more informative statement.

Page 5-26, bullet 4. Surely nature does not provide D-R relationships that follow any simple mathematical formula for any heterogeneous population. The goal is purely empirical -- to find some curve that fits pretty well and is mathematically tractable. For example, a small, unrecognized sub-population of acutely sensitive persons will produce a distinct blip in the toe of the D-R curve such that low-dose risks estimated by a linear model are substantially underestimated. A specific, well-known example is vinyl chloride. Or, an unknown physiologic mechanism may bend the low-dose curve in unexpected ways (e.g., for vinyl chloride). (See Bailar et al., One-Hit Models of Carcinogenesis: Conservative or Not? Risk Analysis 8: 485-497, 1988.)

Page 5-31, line 17. Please expand on "subset".

Section 6. I would like to see some discussion of the level of detail (especially geographic detail) that will be published at the end of the analysis. National only? By state? by county? By SMSA? Other?

**Dr. Michelle Bell**

I appreciate the efforts to define short-term and long-term exposure in this document. This should match definitions used in the ISA. Currently they do not (see page 5-1). Also, short-term exposure is defined in the Scope and Methods document as ">24 hours", which appears to be a typo.

The use of BENMAP is appropriate. The description of BENMAP gives the impression that BENMAP can do things one cannot do without BENMAP. However, every task described could have been done before BENMAP existed and can be done currently outside of BENMAP. The text should accurately reflect BENMAPs abilities, but not overstate them. Be clear that BENMAP is a tool. The benefits of using BENMAP are not the ability to do these tasks, but the ability to do them efficiently and conveniently in a central system with quality control.

The inclusion of health endpoints with supportive evidence for causality but with uncertainty is appropriate as a sensitivity analysis. However, the birth outcomes literature for ozone is quite uncertain, more so than the health impacts of particles on birth outcomes. This particular set of health outcomes may be too uncertain to include.

The fused monitor/model data is appropriate, and is preferable to the use of modeling data alone. This should not supplant the use of monitor data **alone in a separate analysis.**

The analysis using locally-derived estimates needs to address complications such as different studies being used in different cities for the same health endpoint. The local study may not be the best studies. Other issues are different ICD codes for the same health endpoint. Bayesian estimates are far preferable to Maximum Likelihood Estimates. Any adjustment to discount in any way the negative Bayesian estimates needs a detailed justification.

**Dr. David Chock**

The Scope and Methods Plan is quite comprehensive. The EPA should be commended for this thorough effort.

Exposure Analysis:

The Plan intends to cover three or more of the 12 urban areas for exposure analysis. The methodology of Air Pollutants Exposure (APEX) model is sound. The details involved as described in Figure 3-1 of the Plan is quite comprehensive. One concern I have is the input data, especially those from the Consolidated Human Activity Database (CHAD). Many of the database as described in Table 3-2 is quite old, some as old as the 1980s. The RTI NSAS database for eight cities is the most current (2009) but the sample size is relatively small on a per-city basis. Perhaps one can still lean toward enabling the use of these most up-to-date data in the selection of the urban areas.

The list of population groups selected for exposure estimates is adequate, so long as the exposure data are available for each of the selected groups.

The use of a two-dimensional Monte Carlo sampling approach to study both uncertainty and variability of exposure is an excellent idea. The authors point out the differences between model parameter uncertainty and model formulation uncertainty. Indeed, the latter is more critical for model evaluation, but is more difficult to characterize.

In the exposure and health risk assessments dealing with the policy-relevant background (PRB), it would be necessary to use the same metrics as those used in exposure and risk estimates for the present ozone levels. If the annual fourth highest value or the full ozone distribution is used in one, it should be used in the other. (In this connection, there may be some ambiguity for sites that have a limited temporal coverage for the ozone season since the actual high ozone concentrations may start being observed in Spring as the NAAQS starts approaching the PRB levels.) Since GEOS-Chem (and other chemistry transport models) tends to underestimate the upper extreme values, perhaps one can simulate a realistic PRB ozone distribution by adjusting the observed ozone concentration distribution of a remote CASTNET site closest to the area of interest such that the mean of the resulting distribution coincides with the PRB mean concentration estimated by GEOS-Chem.

Health Risk Assessment:

From the description of the controlled exposure studies in the Plan, the model of McDonnell, et al. (2010) for $FEV_1$ seems to cover the full range of ozone exposure concentration. Accordingly, the EPA's proposed approach to estimate the probabilistic lung-function decrement risks is justified.

Section 2.3.2 of the Plans describes the intention to re-evaluate the quadratic rollback approach for the portion of ozone concentration that exceeds the policy-relevant background. This is helpful, especially if the evaluation could involve some stress tests with photochemical modeling, since ozone formation depends on many factors, and the levels and locations of peak ozone can change depending on changes of the precursor emissions and of the relevant atmospheric environment. As extreme weather becomes more commonplace, there is a concern that even the policy-relevant background cannot be assumed to be stable.

It may not be reasonable to apply the quadratic rollback point by point for the portion of ozone concentration exceeding the PRB mean value because doing so will distort the realism of the resulting concentration distribution by suppressing the PRB ozone concentration distribution into a delta function. In reality, as the NAAQS approaches the PRB, the ozone distribution should be more and more PRB-like, with possibly high values in Spring. An ad hoc and fairly realistic but not really satisfactory way out is to use the observed CASTNET ozone distribution adjusted to recover the PRB mean as the fluctuating point-by-point baseline concentration instead of the PRB mean itself for the quadratic rollback, with no rollback assumed when the pre-rollback concentration is less than the same-day PRB-based concentration.

It is reasonable to favor the use of distributed lags over direct multiple lags in short-term risk analyses to reduce multi-collinearity of the lag coefficients.

The ISA does not seem to indicate the presence of thresholds for short-term risks, including lung-function decrements based on control studies. It appears reasonable to assume no threshold provided that the health endpoint risk increment estimates and associated uncertainties are always extrapolated from the actual baseline studies, both upward and downward in exposure concentrations. This should be the case even when alternative, more stringent standards are considered relative to the current standards. The description on page 5-19, lines 20 to 24 needs to be revised so that two $\Delta x$'s, one based on the current standard and one on the alternative standard, need to be determined, both with the same reference point from the baseline studies. Because the model is nonlinear, use of the current standard levels as reference point is inappropriate, especially when the population number is large. For long-term risks, the Plan intends to use a range of thresholds including none. Again, this is reasonable provided that the associated uncertainties are also indicated.

The Plan describes a very comprehensive uncertainty and variability approach in risk assessment. This is commendable. One concern is that there may not be sufficient amount of reliable information to adequately characterize the uncertainty for each of the range of variability to be considered.

In the description of single vs. multi-pollutant models (page 5-13, lines 23 to 28), it should be noted that if the co-pollutants and ozone are all causal variables to the health endpoint, then their inclusion in the model is necessary, and the model based exclusively on ozone would actually be a misfit especially when the ozone correlation with one or more of other causal co-pollutants is large. In fact, multi-pollutant models are a safer choice since non-causal covariates would not impact the coefficients of the causal covariates.

**Dr. Judy Graham**

**General Comments**

1. OAQPS has a stellar history of developing excellent Exposure Assessments (EA) and Health Risk Assessments (HRA). Based on my reading of the plan, this new effort will rise to that historical standard. With appropriate resources, I expect that this effort will even exceed their previous quality. A few specific comments follow, but they do not detract from the general excellence of the plan.
2. I take note that the Plan adheres to most (and perhaps all) of the major recommendations of the NRC on "Models in Environmental Regulatory Decision Making" (2007 http://www.nap.edu/catalog.php?record_id=11972). Chief among them are (1) involving the policy user in setting the goals for the development of a scientifically robust model for the regulatory purpose (in this case a NAAQS), (2) having peer and public reviews of the plan as well as the models during the life cycle of the model, (3)quantifying and communicating uncertainty, (4)evaluating the model throughout its life-cycle, (5) providing leadership for model advancement with an interaction of models and measurements, and (6)using models that are publically available. I recommend that the Plan include a short discussion of the relationship of their proposed use of models to the NRC recommendations. This will add strength to the plan and its eventual results.
3. The situation of resources limiting the EA and HRA may be real, at present, but it is *totally unacceptable*. O3 is one of the very very few pollutants for which ambient exposures cause adverse health effects in a large number of people. Hence, the need for a scientifically robust EA and HRA to enable regulatory decisions to fulfill the requirements of the CAA. Because the NAAQS is already at the third decimal place (0.075 ppm), evaluating its adequacy and the potential need for revision call for a higher degree of accuracy, which in turn requires a thorough analysis. Thus, addressing all high priority EA and HRA needs is essential and must be of highest priority.

**Specific Comments**

1. 1-9 L8. This is the first of several areas where the extent of the effort "will depend on the available resources." As I mentioned above, this is unacceptable. I suggest that OAQPS expand their prioritization here and elsewhere and discuss the value of more completeness to the quality of the NAAQS review. For example, on this page, what is the value of modeling exposure in 3, 6, 9, or 12 urban areas? The Plan discusses good criteria for selection of the urban areas, but are there 3 urban areas that fulfill all these criteria? If not, this would be important to state to clearly identify the need for adequate resources.
2. 1-9 L22. I presume you mean "Quantitative" health endpoints. If so, say so.
3. 3-6 Table 3-1. This lists the microenvironments to be modeled. Which of these scenarios represents "near home"? Also, collectively what percentage of an individual's activities is represented here? For example, is 100% of a child's time accounted for? An outdoor worker's time? Etc.

4. 3-12 L8 This says OAQPS is considering modeling indoor sources of O3. Given the inadequate resources, are there sufficient sources to make this worthwhile? Will you be doing a sensitivity analysis to determine its priority?
5. 3-12 L18. CHAD will be used. It is superb and has been extremely useful for many analyses from the research level to the regulatory level. However, activity patterns are changing, especially for children in this day of sitting in front of the TV or computer, to say nothing of sitting in a chair texting. Thus, please consider the validity of using activity pattern data on children that were conducted prior to 1990, for example. I say "for example" because there might be a more accurate cut point. Since children often fall into the susceptibility class because of their greater amount of time outdoors with exercise, choosing the most appropriate cut point becomes important.
6. 3-16 L20. I strongly recommend incorporating the ATUS data files into the analysis, assuming OAQPS does a professional evaluation of the quality of ATUS (for example, does it have metadata and has it been subjected to extensive QA?). Because of the importance of activity patterns to the EA outcome and the question about the present-day accuracy of older activity pattern surveys, using the ATUS data would be a very high priority.
7. 3-17 L7. Again, the "if sufficient resources are available". Is this possible to do with current databases? If so, do it.
8. 3-19 L18. Are the exposure data and metadata publically available from DEARS. Has the exposure data been subjected to QA? If not, how do you intend to deal with this?
9. 4-1 L23. I agree that selecting FEV1 is a good choice to estimate exposure-response relationships. Some other endpoints do not have as robust exposure-response curves, but they are quite important (e.g., inflammatory markers, airway responsiveness). Given the statements in the text, I am expecting to see a high quality discussion of this that clearly lays out the risks that are likely at low levels, but can't be quantified. I further expect that this discussion will include the concordance of animal toxicology studies since that adds scientific strength to understanding the whole possible array of effects and their severity.
10. 4-4 L6ff. Since the NAAQS rests heavily on analyses of human clinical pulmonary function studies, it is essential to find the resources to see if a computational improvement is possible. It is also important to use *all* of the high quality data available. If not, the quantitative HRA results will be vulnerable to charges of bias from only using the older studies and not incorporating as much science as possible. The presentation clearly described use of the new data (e.g., Kim et al. in press). Therefore, the written plan should be clarified on this point.
11. PRBs are discussed in several places and are defined well. However, it is still not clear how OAQPS will relate the science of biological thresholds/no thresholds and PRB. This is even more difficult because, as described in the plan, backgrounds are increasing due to global factors.

**Dr. Daniel Jacob**

**Air Quality Considerations in the REA document (chapter 2 and chapter 3)**

This chapter of the REA document provides the atmospheric basis for the exposure analyses. I am concerned about the use of 2008-2010 ambient data for the exposure analyses because 2009-2010 are considered to be low-ozone years for reasons having to do with meteorology and possibly the economy. 2006-2008 would be much more representative. In addition, the available PRB calculations from GEOS-Chem are for 2006-2008, and temporal coincidence is very important for sites where the PRB can make a large contribution to total ozone concentrations as in the intermountain West. If EPA decides to keep 2008-2010 as basis for its exposure analyses then GEOS-Chem PRB calculations will be needed for that period. However, a better option is to use 2006-2008.

Also, the EPA needs a strategy for correcting GEOS-Chem biases in PRB estimates. I think that it is useful to distinguish between two types of biases:

- Regional biases, such as in the Southeast US in summer where the model background is too high. A simple correction (and probably good enough) would be to use regionally representative sites (such as CASTNet) and attribute GEOS-Chem biases relative to observations proportionately to PRB and to North American sources. More sophisticated corrections are possible by comparing model and observed frequency distributions of ozone at these sites but they may not be any more accurate.
- High-ozone events in the observations that may be related to PRB and that the model doesn't capture. From my inspection and understanding, I think that these happen only at mountain sites in the West and are associated with stratospheric intrusions or wildfire influences not captured (or excessively diluted) by the model. From my analysis of model vs. observation statistics (and this will be reported in the Zhang et al. [2011] paper describing the GEOS-Chem PRB calculation), the model can properly capture the overall frequency of events > 70 ppb but fails above 75 ppb. Individual inspection of these events in the observations may be necessary to screen for PRB influence.

A few other specific comments:

2-5, line 4: proper reference is Wang et al., AE2009, instead of Bey et al.

2-6, lines 26-28: according to the IPCC AR5 RCP scenarios methane is not projected to further increase in the future. These scenarios may turn out to be wrong, but one cannot just assume that methane will continue to increase.

3-17, lines 23-30: I'm surprised that not more attention is paid to near-roadway exposure. The report states that ozone would be lower because of titration by NO to $NO_2$, which an uneducated reader might assume would reduce exposure, but in fact ppb for ppb $NO_2$ is no better than ozone.

**Dr. Fred J. Miller**

**General Comments**

The document was written with enough detail that a reviewer could see what is planned at the "30,000 foot level". Rather than responding to the list of topics provided by OAQPS for the Exposure Analysis and Health Risk Assessment, I will provide a set of specific comments that collectively cover most of the areas about which OAQPS staff requested that the Ozone Review Committee provide consultative advice. Thus, if I do not cover an area, one can assume that I agree with OAQPS's proposed course of action.

Too many places in the document, the statement is made " we will do this if there are enough resources". Given the uncertainty of EPA budgets, does OAQPS have a strategic priority list that identifies the base things that will be done and then which additional analyses will sequentially be added if the monies are made available? If such a list does not currently exist, one needs to be developed, and a good time to get feedback from CASAC would be when the first external review draft of the health risk assessment is reviewed at a public meeting.

**Specific Comments**

| Page, line | Comment |
|---|---|
| 1-9, 4 | The document states that EPA plans to model population exposures in three or more of the 12 urban areas listed here. This is too vague. If only 3 are done, they should probably represent cities in each of 3 geographically varied regions and the city in the region with the dirtiest air and the highest population used. If 6 were done. One might do the 3 regions together with the cleanest and dirtiest cities. And the list goes on. OAQPS needs to establish criteria for city selection that are based on defined goals and issues. |
| 2-1, 26 | Can a portion of the analyses use a 3-year moving average? Suppose one city is picked and risk estimates are derived using the moving average of $O_3$ levels in that area covering years over at least a 5-year period. This would help establish how dependent the risk analysis estimates are on the air quality data, and it would show if risk has been reduced or increased over time. |
| 3-3 | The figure legend has an oval as representing a simulation step but the figure contains no ovals. |
| 3-7, 24 | Breathing rate is not ventilation rate. Minute volume is the product of breathing rate and tidal volume. Ventilation rate can be minute volume or it can be the inspiratory flow rate. |
| 3-8, 6 | Conflation always has been a problem. One occurrence for 10 people must involve less risk than 10 occurrences for 1 person. Isn't there a way around this? Draw from a sampled pool where each entry has a unique identifier, and then tract the identity of those selected? |
| 3-10, 14 | How is the proximity factor, $P_{FR}$, estimated? |
| 3-11 | In equation 12, one can hardly read the + and – signs. |
| 3-14 | CHAD has only one reference beyond 2005. With the advent of computer |

| | |
|---|---|
| | games, etc., I imagine that activity patterns have changed since most data in CHAD were collected. Thus, it seems critical to get the ATUX data into CHAD or even to replace the current CHAD data. |
| 3-17, 11 | Staff plan to examine averting behavior if resources are available. My view is that it should not be examined because people should not have to change their activity pattern due to pollution – more later. |
| 4-1, 5 | Figure 2 is really Figure 4-1. |
| 4-1, 18 | Where will OAQPS get the data to develop lung function risk estimates for asthmatic school children? |
| 4-2, 4 | To the data being considered for exposure-response relationships, you now also have the data at 60 ppb from Kim et al. (2011). |
| 4-5, 7 | For the piecewise linear fit, why not let the data determine the joining points rather than forcing them because of the data set intervals. This would involve isotonic regression (see Gaines and Rice. Amer. Nat. 135:310-317, 1990). |
| 5-2, 2 | You have enough to do; do not expand to include additional health categories beyond those already identified. The picture is far from clear on birth weight and long-term exposure mortality. |
| 5-6, 4 | What is meant by "First, once we have properly specified the BENMAP software, …."? |
| 5-11, 20 | Long-term exposure mortality seems to be a disconnection. Saying it is a type of respiratory health effect (likely causal) but also a mortality endpoint that is only suggestive of causal is trying to have it both ways. Maybe the CASAC Ozone Review Panel deliberations on these endpoints discussed in the ISA will clarify this. |
| 5-14, 13 | I agree with placing greater weight on using C-R relationships that reflect adjusted single-city estimates obtained form multi-city studies. It represents the best of both worlds. |
| 5-15, 18 | I have a problem with no thresholds. Please clarify how you handle risk below PRB levels. Certainly, there are effective biological thresholds for these endpoints. You are basically making a policy judgment when it should be a scientific one. That being said, it would be a good idea to include a threshold model in a sensitivity analysis to see how much the lower exposure levels contribute to the overall risk. |
| 5-16, 1 | Now the more later. Risk aversion behavior is a negative. Discussions at last week's public meeting clarified what you are attempting to accomplish with risk aversion analyses. It will be important to convey in your findings how risk aversion tends to also negate recommended amounts of exercise for the population, particularly for sensitive individuals. |
| 5-18, 2 | It is tremendously important that you be able to map ICD-9 to ICD-10 codes. Suggest you draw in the main players at CDC that redid the code, and make it their contribution. START NOW even though some at the meeting felt that there is enough lumping of codes that this is not a significant problem. |
| 5-18, 3 | Please clarify cause-specific admissions and ER visits. What about an ER visit or a hospital admission for pneumonia but it is really aspiration pneumonia due to swallowing dysplasia from a stroke? Discussion at the |

| | |
|---|---|
| | meeting indicated that even if such pneumonia incidents are not captured in the death certificates, the bias is present in both the baseline and the epidemiology time series analyses so that this is not a critical factor. However, one has to assume that such occurrences appear equally in all major areas of the country. |
| 5-19, 25 | Don't you have sufficient sample size to use asymptotic variance estimates? |
| 5-22, 16 | Can you fix the values for those model elements that don't have enough data and then incorporate the probabilistic assessment for those elements having sufficient information to assign probabilities? If I go by what is written here, one never would be able to do a probabilistic assessment because a universe of data would be needed. |

**Dr. Armistead (Ted) Russell**

The 2011 Ozone Scope and Methods for Health Risk and Exposure (S&M-Health) is largely derived from the last REA conducted for the prior ozone NAAQS review, which should lead to some efficiencies, e.g., in the application of APEX, though it means it will have some of the weaknesses as well. However, it does appear that they are ready to conduct a more comprehensive uncertainty analysis, which is good. Conducting a thorough uncertainty analysis should be a priority with the infrastructure and appropriate data being built up from the beginning. As noted, there may be limitations, but do not let those interfere with the level of analysis discussed in the S&M plan.

The air quality section notes that their analysis of current conditions will rely predominantly on the data available from AQS, which is appropriate. They note they may use non-AQS data in areas where epidemiologic studies used non-AQS data. This is fine, and little difference is expected for a pollutant such as ozone. The discussion about fusing modeled and observed data sounds like more trouble than it is worth in this case, and may actually introduce issues and require more work to show that the results are better. Similarly, it is not apparent what might be gained by the route explained (not very specifically) at the bottom of 2-2. For one, how do these manipulations conflict with how the epidemiological studies were done, or how APEX would use the data? This needs to be addressed before going down this road.

The section on using a quadratic rollback is not well explained (nor was it last time around). For one, in what independent variable is it quadratic? A diagram would help. Also, they should address how they would modify concentrations that are under the PRB with reasoning.

Use of GEOS-Chem or a similar global model is an appropriate route to take to determine PRB. However, it is imperative that an approach is identified to address known biased results (high or low). This might be a case for data fusion, i.e., to identify the degree to which the base case is biased (using more rural monitors) and how this likely suggests that the PRB should be scaled back/up. I'm of the opinion that the model is more accurately predicting the ozone formed from anthropogenic emissions than getting the transported ozone just right, so that might also suggest scaling back the PRB by the absolute amount that the simulated ozone is above/below the observations.

It was not apparent why they would use GEOS-Chem 2006-2008, but observations from 2008-2010. If GEOS-Chem results are not going to be available, state that specifically, but it would be nice to see what can be done to make the two overlap (even going to 2007-2009 has some attractiveness in terms of a direct comparison over two years, including seeing if the trends agree). Also, the discussion of methane was a bit confusing. What I take from this is that it is not possible to remove North American anthropogenic methane emissions, so there is a need to fudge this because studies shown methane emissions are of some importance in terms of global ozone. It would help if the reason for why methane is being treated the way it is was made clear.

The REA should provide some quantitative information as to the likely change in PRB levels over the next decade and discuss the implications on the potential health (and welfare) benefits over time.

I applaud their continued use of exposure modeling, and that they would appear to be ready to do more uncertainty analysis. This should be a commitment. Only three cities seems to few. Their list of twelve

cities contains appropriate ones, though some prioritization would be good, and I would look first at the cities that have the highest ozone levels, with some consideration as to geographical coverage.
Spell out the name of the Detroit Study.

They should also discuss the potential use of other exposure modeling approaches and why APEX is the current choice over, say, SHEDS.

**Dr. Helen Suh**

The Ozone Health Assessment plan provides a clear and well-thought out outline of the scope and approach that will be used to estimate ozone exposures and health risks under a variety of scenarios. The use of the APEX model to estimate exposures for controlled human exposure studies is appropriate. However, its use may result in ozone exposure estimates that are extremely low, especially for elderly individuals, who spend most of their time indoors where ozone concentrations tend to be low. These low exposure estimates will likely be much lower than the administered exposures in the controlled human exposure studies, raising concerns regarding the applicability of the dose-response relationships from the controlled human exposure studies at these low exposures. If so, alternative or supplemental approaches to estimate ozone exposures for the controlled human exposure studies may be needed. In addition, the Consolidated Human Activity Database will be used to simulate longitudinal activity patterns, which is appropriate given the lack of alternate human activity databases. Since this data base is dated, it may not accurately reflect activity patterns, especially for children and the elderly, contributing what might be a significant amount of uncertainty to the estimates. The magnitude of this contribution should be considered and if possible qualitatively estimated.

**Dr. James Ultman**

My only comment relates to the method by which O3 background levels are taken into account in the risk analysis that uses the McDonnell exposure-response model. As pointed out on page 4-2, this model is based on chamber data in which the pre-to-post change in FEV1 during O3 exposure is corrected for bias by subtracting the pre-to-post change in FEV1 during clean air (i.e., zero background) exposure. According to page 4-3, EPA plans to determine the risk relative to background by using the FEV1 predicted by the exposure-response model to estimate the risk at the O3 exposure condition of interest as well as the risk when the individual breathes air containing PRB O3. The risk relative to PRB will then be computed by the difference between these two separately-computed risks.

Another approach is to predict the FEV1 decrement (corrected by clean air) at the exposure condition of interest, predict the FEV1 decrement at PRB O3 (also corrected by clean air), and take the difference between the two to determine the increased FEV1 decrement relative to PRB. This increased FEV1 decrement could then be used to compute the increased risk of O3 exposure relative to PRB.

Because the computation of risk from a FEV1 decrement is probably a non-linear mathematical process, it is possible that the numerical results from the two methods are substantially different. On the other hand, because the pre-to-post FEV1 decrement in clean air is usually close to that in PRB O3, the two methods may be quite similar even if non-linearites enter the computations. Even so, I believe that the latter method (i.e., computing the increased risk from the increased FEV1 decrement between O3 and at PRB air) is the more logical.

**Comments on *Ozone National Ambient Air Quality Standards: Scope and Methods Plan for Welfare Risk and Exposure Assessment* (April 2011)**

**Dr. John Bailar**

OZONE NATIONAL AIR QUALITY STANDARDS: SCOPE AND METHODS PLAN FOR **WELFARE** RISK AND EXPOSURE ASSESSMENT

I have not reviewed this document is detail, but a survey shows a few points that might be strengthened.

Page 5-3, first bullet. Averages are generally fine when responses are linear in dose or exposure, but may be quite misleading when a dose-response curve (or equivalent) is non-linear. This should be discussed, with some comment on whether average (rather than, say, median, or some percentile) will serve the public best.

Page 5-3, second bullet. Please say more about how the areas will be selected, with special attention to the choice of areas thought to be "representative" in the sense that results can be used elsewhere vs. areas thought to present a special problem (which should not be generalized except perhaps to other areas of the same nature).

Page 5-4, second bullet. The analysis should be explicit in examining the role of co-pollutants as confounders (which could be dealt with by such methods as standardization or stratification) vs. those that may act in synergy with ozone (which cannot be dealt with that way).

The document should say that there are no other known welfare issues, or say what is being omitted.

The dearth of measurements in rural areas, at high altitudes, and elsewhere is a serious impediment to analysis. I would like to see a short section on what the review team sees as the most important areas for research before the next assessment.

**Dr. David Grantz**

**Comments on Welfare Scope and Methods Plan. May 2011.**

Note: Locations in the text are referenced as [chapter-page/line].

## Overview

The approach is reasonable. It is clear that data will not be available to quantify ozone-induced losses in many of the metrics to be evaluated. However, it may not be sufficient to catalog the level of services currently available and thus potentially at risk. More attention should be paid to estimation of loss, even by qualitative means.

In many cases there is a stated intent to evaluate various contrasting methods. In each case it would be useful to describe the criteria upon which the decision would be made. This applies particularly to the various modeling paradigms suggested for air quality interpolation, crop and forest impacts, and crop and forest valuation

## Air Quality

The GIS based national ozone layer is an important goal. The Voronoi nearest neighbor approach may be suitable in the Eastern US, but some reason might be presented to reject the use of a monitor-adjusted CMAQ approach. A uniform paradigm across the country might have certain advantages.

A county scale air quality matrix is adequate in many areas, but in mountainous areas a unit based on an air basin or region (for example in partly mountainous counties of California) might be preferable.

## Ecological Effects of Exposure

The use of NHEERL and NCLAN C-R functions for tree and crop species, derived from OTCs, is appropriate and well justified in the draft ISA. It seems awkward to separate the nearly identical use of NHEERL and NCLAN C-R functions into separate sections of the document. Perhaps the C-R and total productivity risks for crops and trees can be consolidated, followed by use of AGSIM, TREGRO/ZELIG or FASOM in the later analysis of Ecosystem Services. For the NCLAN C-R functions an uncertainty term related to decreasing ozone sensitivity of current cultivars relative to NCLAN era cultivars should be added.

Use of visible injury should be limited to aesthetic values and associated monetization based on willingness to pay. It is increasingly clear that visible injury and other metrics of injury do not correlate well. The National Park study areas are quite appropriate, and represent the best data available.

The ecological metrics to be considered are appropriate, but relatively standard. In the ISA a great deal is made of the potential increase in mean stomatal conductance in forest species and reduced watershed runoff. In some areas, this may have ecological implications to wetlands, riparian species, etc. Furthermore, there may be monetary implications, through irrigation costs (including avoided ground water pumping costs) and lost hydroelectric generation. It should be noted that the increase in mean conductance is probably due reduced stomatal reactivity to closing stimuli, from a partially closed state, rather than actual stomatal opening.

**Ecosystem Services**

The planning document could evaluate the trade-offs involved in using FASOM versus other approaches to crop and forest impacts. This possibility and its likelihood of implementation come up several times. Provided that updated economic data are utilized, the modeling approaches are appropriate.

The ecosystem services concept is very appropriate to this analysis. The list of candidate services is quite complete, and represents a nice target which may or may not be achieved. When monetary values can be assigned this will be useful, but lack of valuation or willingness to pay should not prevent risk analysis. For example, the degraded ability of bees to find flowers of agricultural species for pollination will have monetary consequences that might be possible to estimate. This could be a useful exercise, despite being largely qualitative at this stage.

Continued evaluation of the form of the standard is very important. Repeated assessments have concluded that cumulative ozone exposure is important and high concentrations are particularly significant. Yet it has remained unclear whether the increased complexity of forms such as W126 are sufficiently better at predicting impacts to be worth the increased difficulty of communication to regulated communities.

**Dr. Howard S. Neufeld**

Comments on Ozone National Ambient Air Quality Standards: Scope and Methods Plan for Welfare Risk and Exposure Assessment

The REA seems quite reasonable, if not ambitious, given the amount of variability and uncertainty in the data with which the EPA has to work. I had no major comments on the goals or the structure of the approach for calculating the risk assessment.

One comment concerns the use of FASOM to assess the economic impacts of ozone on natural ecosystems (Section 1.3, page 1-11). A following statement by Staff alludes to using this same model to assess how ozone affects ecosystem services, such as carbon sequestration. Dr. Han Tian, at Auburn University, School of Forestry and Wildlife Sciences, has developed a very sophisticated model called the Dynamic Land Ecosystem Model (DLEM) for determining major biogeochemical cycles as affected by a suite of environmental variables. As stated by Ren et al. (2011): "*DLEM is a process-based model that couples major biogeochemical cycles, hydrological processes and vegetation dynamics to generate daily, spatially explicit estimates of water, carbon ($CO_2$, $CH_4$) and nitrogen fluxes ($N_2O$, $NH_3$, $NH_4$, $NO_x$) and pool sizes (C and N) in terrestrial ecosystems (Tian et al., 2005, 2010a,c; Ren et al., 2007a,b; Liu et al., 2008; see Fig. 1). The DLEM includes five core components: (1) biophysics, (2) plant physiology, (3) soil biogeochemistry, (4) dynamic vegetation, and (5) disturbances, land use and management*."

This appears to be a potentially useful model for EPA use, especially since it has been used previously to estimate ozone effects, both in the Smokies and in China. Here is the citation for the model:

Tian, H.Q., Liu,M.L., Zhang, C., Ren,W., Chen, G.S., Xu, X.F. & Lu, C.Q. (2005) *DLEM – the Dynamic land ecosystem model, user manual*. Ecosystem Science and Regional Analysis Lab, Auburn University, Auburn, AL.

If TREGRO and/or ZELIG are used to model tree and forest responses to ozone, will there be any inputs for rising $CO_2$ and/or elevated temperatures and precipitation? Seeing how rising $CO_2$ moderates ozone responses in the Aspen FACE system it would seem to suggest that some sort of accommodation should be made for this. As for temperature and precipitation, I do not think there are sufficient data at this point in time for the former factor, but there does seem plenty of information that could be used with the latter factor. We still need considerable research on how climate change can alter potential ozone responses. Since rising $CO_2$ generally ameliorates ozone responses, the risk assessments could overestimate future responses to ozone if they ignore this factor.

On page 3-2, in section 3.2.1, Staff list as one goal to estimate the "Percent of trails affected by ozone injury." What exactly does Staff mean by this? Is this simply the relative number of trails where foliar injury was observed? If so, I don't see how relevant that is, and it seems especially trivial when compared to the previous item, which estimates the percent of the vegetation with foliar injury.

In Figure 4.1, what is meant by "Habitat Loss"? Is Staff implying that ozone causes habitat loss? I would not agree with that statement. Perhaps they can clear this up.

Recent papers have documented that farmers are extending their planting seasons in the Midwest due to climate change. When building the estimates for crop losses in section 4.1.3.1, I think Staff needs to take into account the extension of the potential growing season due to climate change now, and it's potential extension in the near future due to rising temperatures. That is, if the growing season is longer, will that raise yields that partially counteract the negative effects of ozone? This may be particularly important if the extended portions of the growing season lie outside the main ozone season.

The rest of the document seems well written, and the risk assessments proposed reasonable.

Below, I document the few typos I found:
Section 1.3, line 26 – the entire "O3" is subscripted, when only the "3" should be.
Section 4.1.2, line 15 – change "is" to "are" since data are plural.
Section 4.1.3.1, line 2 – change "this" to "these" since again, data are plural.

**Dr. Armistead (Ted) Russell**

The S&M-Welfare provides a reasonably thought out preliminary plan for conducting the welfare REA, though it is rather scant on details. Pages 1-8 though 1-11 identify the right type of questions. Many of my comments about the air quality characterization from the S&M-Health carry over here.

I am a bit uncomfortable with how they plan to characterize ozone in more rural areas, at least as currently presented. They note that there are fewer ozone monitors in rural areas, which is true, though ozone levels are somewhat more uniform. To deal with the lack of ozone monitors they discuss using modeled ozone (e.g., using CMAQ) along with observations to develop a fused field. Then, they are looking to use PRB fields developed by GEOS-Chem. At this point you are looking to mesh ozone "data" from three different origins. However, I would suggest that the models are best at getting the ozone production correct more than the background. Thus (as I discussed for the S&M-Health) a rational approach should be developed and tested as to how to modify the GEOS-Chem PRB using current-day observations. (I am also a bit uncomfortable in mixing years.) This could have additional significance here in that the W126 function is non-linear. Thus, staff needs to spend some time thinking about how to develop consistent "current", PRB and alternative ozone fields, and by consistent I mean that appropriately use the observations and model(s) in a way that give results that are sensible in terms of matching current conditions and respond to controls in the expected way, and have an endpoint consistent with what would be expected if North American emissions were removed. It is understood that however this is done there will be some uncertainties, but it needs to be clearly articulated and justified. While it was good to see that they are undertaking an exploratory study of ozone-elevation relationships, how this information is used adds another degree of freedom in merging the various types of information on ozone levels to get the National Ozone Exposure Surface. Again, start considering how all of this might be used, and to a degree have this inform the next ISA in terms of presenting the background (i.e., methods research) that support the likely approaches.

Justifying how you will adjust values below PRB will be important.

The ISA needs to provide more information that will support your analyses, e.g., evaluation of modeled W126 fields and how W126 has trended over the last years.

In regard to what the Welfare REA will look like, I trust some of the same staff that worked on the NOx-SOx secondary REA and PA are working on this one such that many of the lessons learned will get carried over. Many of the issues regarding estimating the impact on ecosystem services and its valuation would carry over. It might be good to having one of the decision similar case study areas to build up a stronger conceptual model of the welfare effects and processes in one or two areas. (I note that GSM is one of the candidates here.) Again, it would be good to provide the appropriate information in the ISA that will be critical to the REA, e.g., in the NOx-SOx review, having more information earlier on about the Ecoregions that were used in the Policy Assessment would have been useful (I realize that the approaches were more in flux, during the NOx-SOx review due to the novelty of that review).

**Dr. Kathleen Weathers**

The approach offered in Scope and Methods Plan for Welfare Risk and Exposure Assessment makes logical sense to me, but, as I've noted in past comments, the devil's in the details. In fact, I think much can be learned from experience and feedback from other panels, for example the NOx SOx secondary effects reviews. The details of that PA, including linking atmospheric inputs (in the absence of direct measures) to multiple process models to estimate welfare effects, appear to be quite similar to what is proposed here. I think that many of the stumbling blocks and challenges will be parallel.

I encourage EPA to take on the task of evaluating and discussing the myriad spatial scales of analysis and mapping that will be used, how they do, or do not match up, and what will be done to resolve critical differences. For example, if the CMAQ model's output (or any other modeling exercise that is used to generate a national ozone surface) is at 12kmx12km spatial resolution, how does that align with vegetation data from National Parks? What kind of spatial variability in exposure can be estimated over the area of, say, Great Smoky Mountain National Park? And, how will recently collected data across elevational gradients and complex terrain be used to verify or extend the models? I applaud the idea and look forward to the rigorous and creative analysis that will underpin the process.

All of the ecological effects analyses require linking models (whether spatial or process) together. It is useful that these models and functions have been vetted, but the challenge of linking them together in credible as well as informative (for standard setting) ways is large, especially given the paucity of data. In fact, model linking is a frontier in earth system science research; the envelope is being pushed in this standard setting process, which can be a good thing for research as well as for policy.

I've listed below a couple of references that may be of interest in regard to approaches (they do not address ozone, perse) for using FIA data, especially, and examining tree species' responses to both pollution and climate (Canham and Thomas and Thomas et al. papers) as well as comparisons of CMAQ output with ground based measurements (e.g., CASTNET; within Pardo et al. paper).

While I think it is interesting to consider the potential effects of ozone on watershed scale hydrology, teasing apart all of the other confounding and controlling variables, especially with so little data available, will be difficult at best.

Canham, C.D. and R.Q. Thomas. 2010. Frequency, not relative abundance, of temperate tree species varies along climate gradients in eastern North America. Ecology 91:3433-3440.
Thomas, R.Q., C.D. Canham, K.C. Weathers and C.L. Goodale. 2010. Increased tree carbon storage in response to nitrogen deposition in the U.S. *Nature Geoscience* 3:13-17.
Pardo, L. H., Fenn, M. E., Goodale, C.L., Geiser, L. H., Driscoll, C.T., Allen E.B., Baron, J., Bobbink, R, Bowman, W.D., Clark, C, Emmett, B., Gilliam, F.S., Greaver, T., Hall, S.J., Lilleskov, E.A., Liu, L., Lynch, J., Nadelhoffer, K., Perakis, S.S., Robin-Abbott, M. J., Stoddard, J., Weathers, K., Dennis, R. L. Effects of nitrogen deposition and empirical nitrogen critical loads for ecoregions of the United States. *In press. Ecological Applications.*

Some specific questions/comments:
Why isn't the national ozone surface modeling included explicitly in the table of uncertainty and variability analyses listed on page 5-1?
What is meant by a non-linear *programming* model?
Acadia National Park might be included as a case study.
I don't understand how trail data could be used.

**Dr. Peter B. Woodbury**

Overall, this is a very good plan -- it is appropriately ambitious in building upon the analyses used in previous staff papers and addressing a wider range of ozone impacts both quantitatively and qualitatively given the available data. The comments below are intended to contribute to further development of this plan.

**Section-page, line numbers**

2-2. Which alternative definitions of the PRB? How will the PRB be estimated? Can a range of plausible PRB surfaces be developed rather than a single one?

2-3, 6. While consistency with the health RA is desirable, there are several reasons why exposure modeling should differ between health and welfare assessments. For vegetation, the exposure metric should be a cumulative one, such as 3-month W126, rather than an 8-h standard. Also, sensitive vegetation occurs in locations with very low human population density, such as Class 1 wilderness areas, and often these locations are far from ozone monitors, requiring a different approach from that used for areas with dense human populations.

2-3, 10. As discussed above, modeling the current standard is not adequate. The metric for vegetation exposure must be cumulative over many months, as discussed in the ISA.

2-3, 19-21. Will analyses include both high and low ozone years? If such variability would affect the results of rollback methods, staff should consider using multiple climate years.

3-1, 11-13. This is an important improvement to use FIA data to quantify the location of different species. However, it should be noted that ozone may be having an impact on the frequency of sensitive populations of some species, even if such effects cannot be modeled quantitatively based on available data.

3-2,13-16. Staff may wish to consider using Acadia National Park, although ozone exposure values are not as high there as some other locations in the USA, there have been exceedences of current standards and there is a body of site-specific data on ozone sensitivity based on open-top chamber experiments and surveys both using native vegetation.

3-3, 9-14. Since urban ozone concentrations may be quite variable over short distances, and be affected by local sources and air flows, it will be important to assure that interpolated ozone surfaces adequately reflect ozone exposures for the urban study areas. It will also be important to consider interacting effects of temperature, drought, and artificial watering on ozone effects on urban vegetation.

4-2. This figure (4.1) is a good start at listing potential ecosystem service impacts, but as mentioned in the draft, a more comprehensive list is needed, for example including climate regulation effects discussed in Chapter 10 of the draft ISA.

4-3. The suggestion to include human recreation frequency data is useful. However, it is also important to assess whether damage is occurring, whether or not a lot of people see it. For damage such as visible symptoms on leaves, this can be viewed as affecting the quality of recreational experiences,

but it is also a direct indication of damage to trees. Other effects of ozone on vegetation may not be visible, but may also be important, such as changes in competitive relationships between species with different sensitivity to ozone.

4-3, 9-16. As mentioned above, it is important to include climate change and climate regulation impacts of ozone (Chapter 10 of the draft ISA). As noted, secondary or "cascading" effects and interactions, such as effects on hydrologic cycles, fire frequency, or pest attacks on vegetation are likely to be more important than direct effects of ozone on vegetation, but are very difficult to measure or to model accurately. I commend the staff for considering such interactions despite the challenges in quantifying effects.

4-3 and 4-4. For cropping seasons, staff should be sure to account for double cropping and winter cropping patterns in different regions.

4-4, 10-18. For economic impacts, it is important to quantify not just global or national impacts on, for example, crop yield, supply and price, but also local impacts on yield. This has not been done adequately in the peer-reviewed literature that I have seen. For example, if soybean yield is decreased in an area of the country with higher ozone exposure, this represents an economic loss to farmers in that region, even if there are economic gains to farmers in other regions that result in only small effects on price or total returns to farmers at the national scale. In such a situation, ozone-induced yield decreases can be viewed as shifting income from high-ozone areas to low-ozone areas, including areas outside of the USA, since agricultural commodities like soybean have global markets. I strongly encourage staff to consider the sum of such local negative impacts (not just net changes in farm income) in addition to impacts on price at the national scale.

4-4, 22-28. The following reports should be useful for assessing impacts of ozone on forests throughout the Southern Appalachian Forest region. A peer-review process was included in the report process. These may be useful for the ISA as well as the welfare and risk assessment. They were previously publically available on the internet, but may not be at this time. I can provide PDF copies of these reports if they might be useful.

Weinstein DA, Woodbury PB, Gollands B, King P, Lepak L, Pendleton D. 2002. Assessment of Effects of Ozone on Forest Resources in the Southern Appalachian Mountains. Report to Southern Appalachian Mountain Initiative.

Southern Appalachian Mountain Initiative. 2002. Chapter 5: Ozone Effects to Forests. In: Southern Appalachian Mountain Initiative Final Report. August, 2002.

4-5, 1-3. Can FASOM really model ozone impacts on different species of trees, and the effects of ozone on mixed stands of trees? Can FASOM model the ecosystem services that are most likely to be affected by ozone? For mixed-species forests, I expect that total biomass growth will not be the most sensitive impact from ozone exposure.

4-6, line 6-15. It would certainly be useful to model effects of ozone on urban trees, and on the services which they provide. However, I have some questions about the capabilities of this model. Can the i-TREE model adequately represent the effects of ozone on specific tree species? Can it adequately account for interactions of temperature and moisture status with ozone that may occur with urban trees? Is decreased growth rate the most important impact of ozone on urban trees?

5-1, Table 5-1. Sensitivity analyses are useful, but quantitative uncertainty analyses are more useful, because they quantify the range of possible effects that may be expected due to the uncertainty in model inputs. Sensitivity analysis, in contrast, examines the sensitive of model results to variation in the inputs, typically increased or decreased by some fixed percentage. Quantitative uncertainty analysis develops probability density functions to represent current knowledge of the input values, and then propagates these functions through the model with appropriate consideration of interactions among variables, to produce probabilistic estimates of impacts. Such probabilistic results can directly indicate the likelihood of much more severe or less severe impacts than suggested by single value estimate. I encourage the staff to consider using quantitative uncertainty analysis in place of sensitivity analysis to address the topics listed in the table, and discussed on later pages. I also encourage staff to use quantitative uncertainty analysis approaches rather than qualitative approaches whenever feasible.

5-2 and subsequent pages. The discussion demonstrates that staff are considering the important distinctions between uncertainty and variability. It is indeed important to try to separate these concepts, and I commend the staff for doing so. However, in practice it is not always feasible to separate them. Additional careful consideration to identifying key sources of variability and uncertainty is warranted (see below). Another important concept is to search for sources of bias in estimating ozone impacts. There will always be uncertainty in modeling impacts of stressors over large regions of space and over long time periods. However, careful quantitative uncertainty analysis can help to reveal potential biases that may occur with single point estimates.

5-4. The list of sources of uncertainty is useful, but it's not clear whether it includes some important gaps in scientific knowledge, including lack of any data on effects of ozone on many plant species, inadequate data on the sensitivity of mature trees compared to seedlings, etc. The list indicates that there is uncertainty about the policy-relevant background level for individual locations, there is also uncertainty about the level at regional to continental scales, since it can only be estimated using air quality models. Consideration of such issues, particularly if they may cause bias in point estimates is warranted.