Review and Critique of the U. S. Environmental Protection Agency's First External Review Draft of the "Integrated Science Assessment for Ozone and Related Photochemical Oxidants"

By
Jon M. Heuss
George T. Wolff
Air Improvement Resource, Inc.

Prepared for
The Alliance of Automobile Manufacturers

May 5, 2011

Introduction

The U. S Environmental Protection Agency (EPA) initiated the next review of the National Ambient Air Quality Standards (NAAQS) for ozone (O_3) with the issuance of the first external review draft of the Integrated Science Assessment for Ozone and Related Photochemical Oxidants[1] (ISA) in March 2011. Air Improvement Resource, Inc. (AIR) reviewed the draft with a focus on the portions of the ISA that are important to providing the Administrator with the most relevant science with which to judge the health effects of ozone and establish a primary ozone standard which will protect the public health with an adequate margin of safety. AIR and the Alliance of Automobile Manufacturers (Alliance) participated in the previous review of the ozone standard which resulted in the 8-hour standard being set at 0.075 ppm.[2] The Alliance also participated[3] in the re-consideration of the ozone standard that was initiated by Administrator Jackson in January 2010.[4] The following comments focus on the human clinical studies of ozone effects and their interpretation in terms of the public health and the epidemiological studies of associations of ozone with health endpoints and their interpretation in terms of public health. Although these comments do not cover policy relevant background, the

---

[1] U. S. Environmental Protection Agency, First external review draft of the Integrated Science Assessment for Ozone and Related Photochemical Oxidants, EPA/600/R-10/076a March 2011.
[2] Comments of the Alliance of Automobile Manufacturers on EPA's Proposal to Revise National Ambient Air Quality Standards for Ozone, 72 Fed. Reg. 37,818 (July 11, 2007), dated October 9, 2007.
[3] Comments of the Alliance of Automobile Manufacturers on EPA's Proposal to Revise National Ambient Air Quality Standards for Ozone, 75 Fed. Reg. 2992 (January 19, 2010), dated March 22, 2010.
[4] 75 Fed. Reg. 2992, January 19, 2010.

Alliance is concerned that EPA underestimates policy relevant background. Details are provided in previous submissions (Alliance October 9, 2007 and March 22, 2010 comments).

Before discussing the evidence for individual categories of effects, it is appropriate to make some general comments on the organization of the document. Chapter 6 reviews, summarizes, and integrates the evidence for various health outcomes related to short-term ozone exposures. As such, it is the main place in the document where ozone health effects are discussed and will be the focus of AIR comments. However, the information in other Chapters is important as it illuminates and informs the discussion and integration in Chapter 6. Therefore, we will provide comments on other chapters as needed to aid in the integrative discussion.

Chapter 6 is organized by major health effect categories (e.g., respiratory, cardiovascular, mortality). Within each category, the evidence is organized by health endpoint (e.g., lung function, pulmonary inflammation) and then by specific scientific discipline (e.g., controlled human exposure, epidemiology, and toxicology). Each major section (e.g., respiratory, cardiovascular, mortality) concludes with an integrated summary of the findings and a conclusion regarding causality. The determination of causality is made for each broad health effect category, such as respiratory effects, for example. While the overall organization is reasonable, making causality determinations for such broad categories is misleading. For example, the evidence of causality for the various respiratory endpoints varies dramatically. To lump the evidence together and draw a conclusion regarding causality for such a large category is scientifically unsound. Instead, separate determinations should be made for sub-categories of effect.

I. Human Clinical Studies and Their Interpretation

As indicated in the ISA,[5] the controlled human exposure studies provide a strong and consistent body of information on the dose-response of effects of 1-to-3 hour and 8-hour exposures to ozone. The first effects - transient, reversible FEV1 decrements - are the body's reflexive reaction to the presence of an irritant gas unrelated to sensations of discomfort. Such effects occur after exposures to 0.08 ppm for 6 to 8 hours when the subjects are exercising at a rate that would be considered strenuous when carried out intermittently for an eight-hour period. Whether such effects occur at 0.06 or 0.07 ppm has been highly controversial since the answer depends on how the baseline is evaluated, how the precision of the test is considered, how the day-to-day variability of a subject is evaluated, and how the data is statistically analyzed. During the last review, the Adams 2006 study was the only study available at concentrations below 0.08 ppm. The Schelegle et al., 2009 and Kim et al., 2011 studies are now also available.[6] As noted in the ISA these studies now all indicate very small group mean changes in FEV1 at 0.06 ppm. There is also evidence from these studies that a small portion of the subjects had FEV1 changes greater than 10 %.

[5] ISA, supra note 1, at page 6-2.
[6] Literature referred to by author and date without a footnote are references included in the ISA.

The important question is not whether these small changes are statistically significant; the important question is their medical or public health significance. The ISA does not adequately lay the groundwork for answering this question. The ISA refers to two publications regarding guidelines for determining clinically meaningful FEV1 changes. These two references discuss the use of lung function testing to evaluate various obstructive and restrictive disease states that result in changes in lung function. For example, the Pellegrino et al., 2005 review discusses lung function changes as they relate to progressing disease or the response of disease states to therapy. Pellegrino et al. do not discuss the clinical significance of the kind of transient, reversible changes caused by ozone. They do note, however, that statistical significance and clinical significance do not follow one another. They point out that two lung function measurements that are statistically indistinguishable may provide reassurance in a patient receiving therapy for a disease that is otherwise rapidly progressive. They note that the same tests may be very disappointing if one is treating a disorder that is expected to improve dramatically with the therapy prescribed. They also point out that a statistically significant change may be of no clinical importance to the patient.

The more relevant American Thoracic Society guidelines, that are not referenced in the ISA, are guidelines regarding what constitutes adverse air pollution effects. The 1999 Guidelines indicate:[7]

> The committee recommends that a small, transient loss of lung function, by itself should not automatically be designated as adverse. In drawing the distinction between adverse and nonadverse reversible effects, this committee recommended that reversible loss of lung function in combination with the presence of symptoms should be considered adverse.

Therefore, the ISA should discuss the symptom results for the human clinical studies along with the FEV1 results to provide appropriate information to the reader.[8] In the Adams, 2006 study, the total mean symptom scores were only 2-4 units at 0.04 and 0.06 ppm out of a possible total score of 160. Adams indicated that the differences in the symptoms between the 0.04 and 0.06 ppm exposures and the filtered air control were not statistically significant. Kim et al., 2011 also indicate that the symptom scores were not different between ozone and clean air. Schelegle et al., 2009 indicate that the symptom scores were increased at 0.07 and 0.08 ppm but not at 0.06 ppm.

Thus, according to the ATS guidelines, the functional changes at 0.06 ppm would not be considered as adverse. The ISA should expand on the clinical and public health relevance of the functional effects. The basic nature and extent of functional effects has not changed since the 1997 and 2008 reviews. There is now data between 0 and 0.08 ppm, but the assumption made in 1997 was that functional effects, albeit small, do occur

---

[7] What Constitutes an Adverse Health Effect of Air Pollution? Official Statement of the American Thoracic Society Adopted by the ATS Board of Directors, July 1999, Am. J. Respir. Crit. Care Med., **161**, 665-673, 2000.

[8] In the draft ISA respiratory symptoms are discussed as a separate category of effects in Section 6.4, but only epidemiological studies are discussed at that point.

below 0.08 ppm. In the 1997 review, single incidences of the effects at 0.08 ppm (for either healthy or asthmatic subjects) were not considered to be adverse by CASAC and EPA staff. Nothing in the body of controlled studies has changed to alter that view. If anything, the growing evidence that the functional effects are caused by activation of neural reflexes, as discussed in greater detail in Section 5.2.3 of the Dosimetry chapter, should reduce the concern over isolated transient, reversible lung function decrements.

Since the first effects on the performance of lung function tests occur at 0.50 ppm in sedentary individuals, the vast bulk of personal exposures of either the general population or the susceptible population are far below the thresholds for the first effects identified in controlled studies. For example, the data on indoor/outdoor ratios and personal exposures in Section 4.3 clearly show that personal exposures are only a fraction of the levels measured at ambient monitors. Typically, personal exposures average a quarter or less of the ambient measurements, even for school children that spend an average of two hours per day outdoors. Even for a group of camp counselors, the personal exposures averaged less than half of the ambient measurements. The ISA concludes that "Another important finding is that the magnitude of personal exposures is smaller than concentrations reported at fixed-site monitors due to time spent indoors and the low indoor penetration of O_3."[9] This important finding needs to be included in the integrative discussion in Chapter 6. It provides a large margin of safety from the first effects identified in controlled human studies for the population as they go about their daily activities.

Other effects on the respiratory system occur at higher concentrations as ozone triggers other responses in the body's defense mechanisms. The presence of inflammatory markers has been studied in many human clinical studies. The ISA refers to a meta-analysis of 21 studies (Mudway and Kelly, 2004) which showed that neutrophil influx in healthy subjects is associated with total ozone dose (i.e., the product of ozone concentration, exposure duration, and ventilation rate). The studies included in Mudway and Kelly evaluated inflammatory markers in exposures that ranged from 0.08 to 0.60 ppm ozone with varying degrees of exercise and durations from 1 to 6.6 hours. The ISA notes that the presence of neutrophils in the lung has long been accepted as a hallmark of inflammation and is an important indicator that ozone causes inflammation in the lungs. Neutrophilic inflammation of tissues indicates activation of the innate immune system and which is normally followed by processes that clear the evidence of acute inflammation.

The immune system responses noted by EPA as the first indications of "inflammation" are physiological processes that occur in all living organisms under the stimuli of daily life. The first reported changes (that occur in humans with heavy exercise after 1- to 3-hours above a threshold of 0.18 to 0.20 ppm) are small and reversible and well within the range of physiological variability. They fall into the category of biochemical markers that the American Thoracic Society indicates do not necessarily imply adversity. The 2000 review by Mudway and Kelly[10] notes that for neutrophils transiting into the lung --

[9] ISA, supra note 1, at page 6-14.

[10] I. Mudway and F. Kelly, *Ozone and the Lung: A Sensitive Issue*, Mol. Aspect. Med., 21, 1-48

one of the earliest of these responses -- it is not clear if the response should be considered beneficial (functioning to clear necrotic cells) or detrimental (leading to an active inflammation with tissue injury). The 2006 Criteria Document noted that generally, "the initiation of inflammation is an important component of the defense process; however, its persistence and/or its repeated occurrence can result in adverse health effects." Since the threshold for even the first indications of an inflammatory response is as high or higher than that for the reflexive FEV1 response, no adverse health effects are to be expected at the typical current ozone exposures shown in Chapter 3 of the ISA. For example, the threshold for inflammatory changes in the Mudway and Kelly 2004 meta-analysis at a 1-hour ozone concentration of 0.12 ppm is a ventilation rate greater than 10 times the resting rate. The threshold at an 8-hour ozone concentration of 0.08 ppm is a ventilation rate greater than 3 times the resting rate.

As the ISA discusses the implications of these clinical findings, the findings from toxicology and epidemiology are discussed. However, two other factors need to be included in the mix of considerations. The first is a more detailed discussion of the public health implications of the FEV1 and other changes. The second is consideration of real world exposures and how often such effects may occur in human populations. We discuss the implications from epidemiology in the next main section. The other three factors are discussed in the following.

First, the toxicological findings need to be interpreted in relation to the body's natural defense mechanisms. Toxicological findings for ozone are discussed both in Chapter 5 Dosimetry and in sections of Chapter 6. Even though there are many references in these sections to the body's natural defenses against the presence of an irritating and oxidizing gas, there is insufficient consideration of the issue of dose-response. The discussion of mechanisms of action does not include a discussion of the doses that are required to elicit the various effects and pathways. The question of plausibility involves two factors, the kinds of effects an agent may cause and the levels of the agent that are necessary to cause the effects. The question of dose-plausibility needs to be fully discussed in the ISA. With regard to this issue, Oberdorster et al., 2005[11] make several important points. First, a careful evaluation of exposure–dose–response relationships is critical to the toxicologic assessment of an agent. Second, although high dose studies may be used in a first proof-of-principle approach, it is mandatory to follow up and validate results using lower concentrations resembling realistic exposures. Third, the 500-year old phrase "the dose makes the poison" can also be paraphrased as "the dose makes the mechanism." The mechanistic pathways that operate at low realistic doses are likely to be different from those operating at very high doses when the cell's or organism's defenses are overwhelmed.

Second, the public health implications of transient, reversible FEV1 changes need to be considered carefully in the integrative synthesis. Public comments during the previous review have established that a variety of environmental exposures and stresses to which many people, including children, are routinely exposed produce changes in pulmonary

---

(2000).

[11] G. Oberdorster, et al., Environ. Health Perspect.,113:823–839 (2005).

function tests of the same magnitude as observed from ozone below 0.08 ppm. Many children experience minor exposures and stresses that result in transient FEV1 decreases of 10-15%. Further, in some cases, these routine changes or effects occur consistently rather than infrequently, i.e. once per year.[12] As noted above, an important question in interpreting the ozone clinical studies is not whether there are small changes in the performance of lung function tests, but what these small changes mean for public health given the involuntary reflex mechanism causing the changes. The small group mean changes that are at issue are well within the range of normal measurement variability and even the responders/outliers in the data are no more than the 10 to 20% changes which are considered moderate and potentially of health concern in the 2007 Staff Paper.

Another consideration is whether persons with asthma, the elderly, and particularly children, are more sensitive and experience larger decrements in lung function due to O_3 exposure than do healthy volunteers. The ISA summarizes the available data, noting that children, adolescents, and young adults (<18 years of age) appear, on average, to have nearly equivalent spirometric responses to O_3, but have greater responses than middle-aged and older adults when exposed to comparable O_3 doses.[13] The ISA goes on to indicate that symptomatic responses to O_3 exposure, however, appear to increase with age until early adulthood and then gradually decrease with increasing age. This was the state of the evidence in 1996 and is still the state of the evidence in 2011. Noting that the ISA also indicates that asthmatics are at least as sensitive as healthy subjects, studies of healthy young adults are likely the most sensitive means to evaluate lung function test decrements with symptoms, the physiological changes ATS views as adverse.

In the 1997 review, single incidences of the FEV1 effects at 0.08 ppm (for either healthy or asthmatic subjects) were not considered to be adverse by CASAC and EPA staff. The 1996 Staff Paper included extensive discussion of how to interpret the clinical results in terms of public health.[14] Large functional changes, > 20 % FEV1 decrements, and severe symptomatic responses were indicated as clearly adverse. Moderate functional changes and symptoms were discussed in relation to interference with normal activity for both healthy and asthmatic individuals. For asthmatics, the Agency and CASAC concluded that moderate responses, when repeated, should be considered adverse. After considerable discussion there was consensus on CASAC that single, acute moderate functional responses should <u>not</u> be considered adverse for healthy individuals. Rather the staff indicated that the number of exposures resulting in moderate responses should be considered a factor in determining adversity for healthy individuals. In addition, the category of moderate functional changes without symptoms or with minimal symptoms was not specifically addressed in previous reviews.

---

[12] NAM, Attachment 2, Critical Review of the Health Data in EPA's 2008 Proposed NAAQS for Ozone Cited in the Current 2010 Proposed Rulemaking, Docket ID No. EPA–HQ–OAR–2005–0172-12439.4, at pages 1 to 7.

[13] ISA, supra note 1, at page 6-13.

[14] U. S. Environmental Protection Agency, Review of the National Ambient Air Quality Standards for Ozone: Assessment of the Scientific and Technical Information, OAQPS Staff Paper, EPA-452/R-96-007, June 1996, at pages 62 to 72.

As CASAC has indicated, for ethical reasons, controlled exposure studies involve effects that are relatively mild and reversible, including changes in pulmonary function and increased evidence of inflammatory changes. Controlled studies of asthmatics and Chronic Obstructive Pulmonary Disease (COPD) patients have been conducted with intermittent exercise at substantially higher ozone exposures than the current standard, resulting in group-mean FEV1 decrements as high as 20 to 25 %, suggesting that such effects are relatively mild with regard to clinical or public health significance.[15]

The third factor that needs to be considered is exposure – the extent to which human populations as they go about their daily activities experience exposure that will lead to the effects identified in the human clinical studies. This factor is discussed in some detail in the next section.

Factors Which Influence the Risk as Estimated from Human Clinical Studies

The exposure/exertion protocols used in the human clinical studies do not mimic typical human behavior so the results must be mapped onto realistic exposure/activity scenarios to determine the risk. The chance of experiencing an exposure of concern requires that a person be outside at the location of the high ozone, at the time of peak ozone and exercising heavily. In both the 1997 Review and in the 2008 Review, the Administrator relied on a Risk Assessment developed by the Agency to evaluate the risk of lung function decrements. Because of the central importance of the model and assumptions that go into the Risk Assessment, the ISA should evaluate the scientific basis for the Agency modeling in detail. The draft ISA does not do this. Instead the draft provides a short summary of the model in Section 4.4.2 that indicates that the model works rather well.

During the 2008 review, the Alliance provided comments on at least five ways in which the clinical exposure and risk assessment was biased, overstating the exposures and risks that would accompany any of the alternative standards under consideration.[16] Specifically, the Alliance documented three ways that the APEX model used in the risk assessment overestimated the number of occurrences of elevated 8-hour ozone exposures during strenuous work or play and two ways in which the Risk Assessment (RA) overestimated the benefits of alternative standards. The three ways APEX over-estimated the number of occurrences are discussed in the following.

First, the analysis did not take into account that human ozone exposures near a monitor are lower than the monitor measures. The 2006 CD acknowledged that ozone exposure is lower at "person" height compared to "measurement" height (3-15 meters). For example, Wisbeth et al., 1996[17] measured the increment between ozone at 2 and 10 meters and reported an average 13 percent difference. In addition to the height differential, ozone

[15] See Table AX6-3 in Vol. II of 2006 Ozone CD.
[16] Alliance comments, supra note 2, at pages 13 to 17.
[17] A. Wisbeth, G. Meiners, T. Johnson, and W. Ollison, *Effect of monitor probe height on measured ozone concentration*, Paper No. 96-RA111.02, presented at the 89th Annual Meeting of the Air & Waste Management Association, Nashville, TN, June 1996.

monitors are also sited in open areas removed from sources so as to capture the highest ozone concentrations expected in an area. Since downwind sites are usually the design value sites, they will dominate the upper tail of the ozone distribution and yet may not reflect the overall outdoor exposures in the vicinity of the site. If people spend time outdoors in closer proximity to streets or in areas with more surface area (buildings, etc.) to quench ozone, their exposures will be below that measured at the monitor. The APEX model assumes that whatever ozone is interpolated from the monitor measurement is the actual ozone exposure in the outdoors microenvironment. The Langstaff Memorandum acknowledged the issue of vertical variation in ozone but indicated that the Agency does not plan to address it due to a lack of data. This vertical difference was corrected in the vegetation risk assessment in the previous review but not in the human risk assessment. In the vegetation risk, the metric summing concentrations of 0.06 ppm and higher was halved with a 10 percent vertical correction. By analogy, a vertical correction in the human risk assessment would likely halve the number of human exposures of concern at ground level. Because this effect would correct a bias in the exposure calculations, it is particularly important that the ISA include a discussion of the difference between ozone at person height and at measurement height.

Second, APEX overestimates the exposures by the way the clinical results are mapped into the model. The upper tail of the distribution of breathing rates -- equivalent ventilation rates (EVR) -- is particularly important because it is a critical factor in determining the number of exposures of concern. APEX assumes that all exposures between 13 and 27 EVR respond as though they were at 20 EVR, substantially overestimating the risk since there are many more 8-hour occurrences of ozone at 13 EVR than at 20 and many more at 20 EVR than at 27. This results in an overestimation of the number of exposures of concern. The 1996 CD showed that 1- to 2-hour ozone responses vary with EVR and Ollison has shown that the 8-hour responses in the Folinsbee/Horstman data do too.[18] Thus, the dose-response curve varies with EVR. The distribution of exposures also varies with EVR. We expect many more 8-hour occurrences of ozone exposures at 13 EVR than at 20 and many more at 20 EVR than at 27. Thus, the approximation that all exposures between 13 and 27 EVR respond as though they were at 20 EVR substantially overestimates the risk.

Since the model calculates EVR to a much finer scale, the distribution of EVR should be presented in the risk assessment and dose-response curves should be developed as a function of EVR to avoid the overestimation of risk. At a minimum, this issue should be evaluated in a sensitivity analysis. The Langstaff Memorandum indicates that the EVR cutpoints input to APEX are selected for the risk calculations and are not considered as uncertain for the purposes of the uncertainty analysis.[19] However, it would be simple to break the wide EVR range used for the risk calculations into a number of smaller bins to test the hypothesis that the use of such a wide EVR range biases the overall result. In addition, the American Petroleum Institute (API) has provided EPA with a dynamic statistical model that can be implemented in APEX and provide more highly refined

[18] American Petroleum Institute Comments on EPA's Dec. 13, 1996 Proposed Ozone Rule, Appendix A.
[19] Draft Langstaff Memorandum at 43.

FEV_1 decrements.[20] The sensitivity of APEX predictions to the use of the API model should be evaluated by EPA and the results discussed in the ISA.

Third, the model predicts more elevated ventilation rate occurrences than observed in real world data. Langstaff acknowledged that the "values produced by the ventilation rate algorithm may exhibit an excessive degree of variability."[21] The final sensitivity analysis for APEX includes a comparison of predicted ventilation rates with mean values in the literature, but the upper tails of the distribution which impact the risk estimates were not compared.[22] This is an important oversight because the upper percentiles of ventilation rate are responsible for the exposures that cause the perceived risk. In the comparison of the APEX modeled values with the measured ventilation rates from Brochu et al., 2006,[23] the model over-predicted mean daily ventilation rates for persons below age 11 and over age 40. More importantly, the model had a much higher standard deviation at all ages. This suggests that the upper percentiles of ventilation rates in the model are substantially above those measured in a database of over 30,000 person-days from a cohort of over 2200 free-living individuals between the ages of 3 and 96. The following Figure 1 shows that the APEX model EPA used in the prior risk assessment significantly overestimates the breathing rates of male children, particularly for the upper tails of the distribution that are responsible for the exposures of concern evaluated by the Agency. The data underlying these distributions (means and standard deviations) come from Table 25 in the 2007 Langstaff Memorandum on uncertainty in the exposure model. In fact, of the 16 comparisons in Table 25, for eight age groupings each of males and females, 15 had substantially higher modeled ventilation rates compared to the data reported by Brochu et al., 2006 at the upper end of the distribution.

---

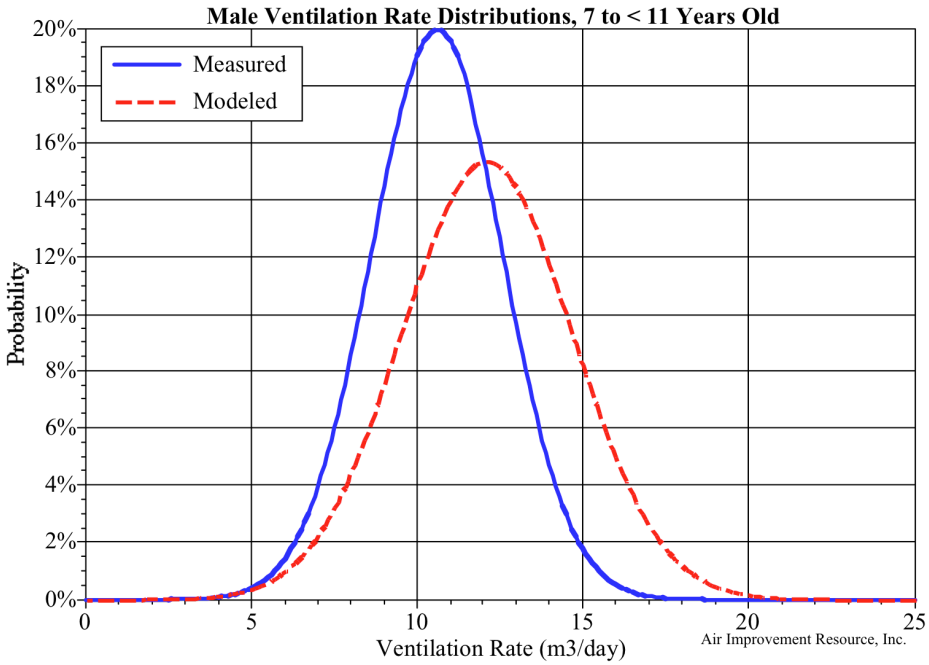
[20] Memorandum from Ted Johnson, TRJ Environmental, to John Langstaff and Harvey Richmond, U.S. EPA, *An algorithm for APEX that estimates one-minute ozone-induced FEV1 decrements*, (July 31, 2006) (included as attachment to Memorandum from Will Ollison, API Energy, to Karen Martin, U.S. EPA (Aug. 29, 2006) (EPA-HQ-OAR-2005-0172-0037.1)).

[21] Draft Langstaff Memorandum at 42.

[22] J. Langstaff, Technical Memorandum, *Analysis of Uncertainty in Ozone Population Exposure Modeling*, Jan. 31, 2007 at 52 (EPA-HQ-OAR-2005-0172-0174).

[23] P. Brochu, J. Ducre-Robitaille, and J. Brodeur, Physiological daily inhalation rates for free-living individuals aged 2.6 months to 96 years based on doubly labeled water measurements: comparison with time-activity- ventilation and metabolic energy conversion estimates, Int. J. Hum. Ecol. Risk. Asses., 12, 736-761 (2006).

Figure 1: Comparison of Measured *vs.* Modeled Daily Ventilation Rates for 7- to 10-Year Old Boys



The 1997 EPA analysis had also over-estimated the number of high ventilation rates in the population. That analysis used an algorithm to assign ventilation rates based on individuals who exercised regularly and were motivated to reach a high ventilation rate. As a result, the 1996 Staff Paper acknowledged that the analysis allowed more high ventilation rates (hence greater risk) than would actually occur in the populations of interest -- outdoor workers, outdoor children, etc. Because of the importance of the methodology for assigning ventilation rates to the estimated risk, the ISA should include a detailed discussion of the methodology and data involved.

II. Epidemiological Studies and Their Interpretation

A. General Comments

In contrast to the science regarding human clinical effects which, while refined, has not changed substantively since the 1997 review, the available epidemiological evidence has increased dramatically. However, along with an outpouring of studies has come increased understanding of the limitations of the epidemiological evidence. For example, publication bias is now known to exaggerate the apparent strength and consistency of association. Limitations due to issues of model selection add substantially to the uncertainty. There is substantial evidence that stochastic variability adds substantially to the uncertainty. In addition, the issue of confounding raises the possibility that a positive association for ozone or any other pollutant in a single-pollutant model may be an indicator of some other pollutant rather than evidence of an independent effect of that pollutant. While one possibility is that ozone may be an indicator of the mix of photochemical oxidants, other possibilities also need to be acknowledged.

The ISA does not adequately discuss these methodological limitations and concerns. It should clearly state that air pollution time-series epidemiology studies suffer from problems associated with publication bias, model uncertainty, model selection issues, lack of adequate control for confounding variables such as other pollutants and weather, and exposure misclassification arising out of the poor correlation between ambient monitors and personal exposure and consider these limitations in the integrative discussions. In a June 2006 letter to the Administrator, CASAC confirmed this view in evaluating mortality time-series studies, noting that "[b]ecause results of time-series studies implicate all of the criteria pollutants, findings of mortality time-series studies do not seem to allow us to confidently attribute observed effects specifically to individual pollutants."[24]

There are severe limitations to the use of epidemiology to try and tease out interactions and to evaluate causality. A meta-analysis by Steib et al.[25] evaluated 109 acute mortality studies from around the world. They report that there are positive associations with mortality (with a wide range in the individual cities) for all the major pollutants in single pollutant models and that for each, when other pollutants are included, the association with the first pollutant, on average, is decreased. In fact, the patterns in single-pollutant epidemiological studies were very similar for all the criteria pollutants.

The studies evaluated by Steib et al. are all subject to publication bias. To avoid publication bias that would inflate the apparent association, investigators have carried out large multi-city analyses. In fact, the patterns in single-pollutant associations in multi-city epidemiological studies are also very similar for all the criteria pollutants. The individual-city associations in large multi-city studies also cover a biologically implausible wide range from strongly negative to strongly positive, a finding which is readily apparent but seldom discussed.

There is also strong evidence for unrecognized stochastic variability in associations within a given city. Ito[26] re-analyzed the 1220 separate air pollution mortality and morbidity associations that were included in the original Lippmann et al. 2000 HEI study of Detroit. As shown in Ito's Figure 2 below, there was a wide range of negative and positive risks in Detroit when all pollutants, lags, and endpoints were considered. Ito showed in separate figures that the wide range of associations occurred for each pollutant. Although the focus in the original Lippmann study, as it is in almost all the published literature, was on the positive associations, Ito's plot shows that there are many negative associations in the data. Although there may be somewhat more positive associations than negative associations, there is so much noise or variability in the data, that

---

[24] R. Henderson, CASAC Letter, EPA-CASAC-06-07, June 5, 2006 at page 3.

[25] D. Steib, S. Judek, and R. Burnett, "Meta-analysis of time series studies of air pollution and mortality: Effects of gases and particles and the influence of cause of death, age, and season," *J. Air & Waste Manage. Assoc*., **52**, 470-484 (2002) and Stieb et al., *J. Air & Waste Management Association*, **53**, 258- 261, 2003.

[26] K. Ito, pages 143-156 in HEI Special Report: Revised Analyses of Time-Series Studies of Air Pollution and Health, May 5, 2003.

identifying which positive associations may be real health effects and which are not is beyond the capability of current methods.

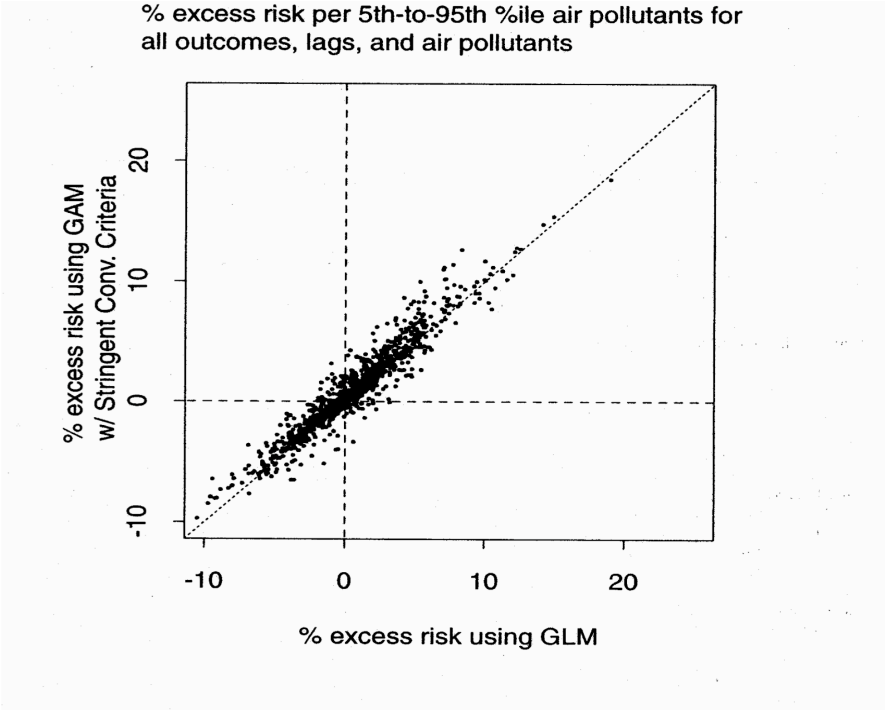


Figure 2 - Comparison of air pollution risks calculated with the General Additive Model (GAM) versus the General Linear Model (GLM) from Ito 2003.

Although EPA is planning on evaluating pollutants in a multipollutant context in the future, the ozone ISA must start that process now by acknowledging and discussing the full range of results for the many studies that evaluated multiple pollutants. Since the 2004 National Research Council report on air quality management in the U. S. recommended that EPA address multiple pollutants in the NAAQS review and standard setting process, there have been several papers/reviews discussing how to do this.[27] At a February 2011 EPA/HEI Workshop on Multipollutant Risk Assessment, EPA staff indicated that the Agency will develop a Multipollutant Science Assessment (MSA) in parallel with the current ISAs for individual pollutants. EPA will be developing the MSA over the next several years and plans to have the MSA inform the NAAQS decisions for single pollutants in coming years.

---

[27] Dominici et al. (2010). Protecting human health from air pollution: Shifting from a single-pollutant to a multipollutant approach. *Epidemiology*, 21, 187-194; Greenbaum and Shaikh (2010). First steps toward multipollutant science for air quality decisions. *Epidemiology*, 21, 195-197; Hidy and Pennell (2010). Multipollutant air quality management. *J Air Waste Manage Assoc*, 60, 645-674; Mauderly et al. (2010). Is the air pollution health research community prepared to support a multipollutant air quality management framework? *Inhal Toxicol*, 22(S1), 1-19; Vedal and Kaufman (2011). What does multi-pollutant air pollution research mean? *Am J Respir Care Med*, 183, 4-6; National Research Council (2004). *Air Quality Management in the United States*. National Academies Press Washington DC.

As each criteria pollutant has been reviewed, EPA is claiming or has claimed independent respiratory effects that are considered as causal or likely causal for four different criteria pollutants and is using selected single-city single-pollutant associations for that pollutant to set the NAAQS. In the recently completed NO_2 review, EPA used a cluster of five selected single-pollutant NO_2 associations that were cited as being "positive, and often statistically significant" to establish the level of the new 1-hour NO_2 standard.[28] The Agency also indicated that single-pollutant NO_2 associations were generally robust to the inclusion of other pollutants. In the recently completed SO_2 review, EPA cited ten studies that reported mostly positive and sometimes statistically significant single-pollutant SO_2 associations and then used SO_2 associations in three cities that remained statistically significant in multi-pollutant models with PM to set the level for the new 1-hour SO_2 standard.[29] In the recent Agency proposal to re-visit the ozone NAAQS, EPA cites epidemiological evidence as a main reason to support the low end of the proposed range for a revised primary standard.[30] In the draft PM Policy Assessment, selected single-pollutant individual city associations are being used to evaluate a range for a potential 98th percentile PM_{10} standard. Only in the case of the recent proposal to keep the current CO standards, has the Administrator discounted the epidemiological associations with CO as possibly acting as a surrogate for other pollutants.

In recognition that the Agency is moving toward a multipollutant approach, and in preparation for that approach, the ozone ISA requires major revision. Revisions are necessary in the way the individual epidemiological studies are presented and discussed in the text, in the way the data are summarized in Tables and Figures, and in the way the material is evaluated and integrated together. In addition, there need to be revisions to the earlier chapters to provide an accurate context for the integrative discussion.

Revisions needed in the main text The current text focuses on selected ozone associations and presents results for multi-pollutant models only to the extent they change the ozone associations. Since many of the studies cited evaluated multiple pollutants in single pollutant models, the overall pattern of associations for the other pollutants needs to be discussed. If the patterns of association are similar for other pollutants for a specific endpoint, that should be brought out in the text. Where available, data on spatial and temporal variations in associations should be discussed. For example, when multi-city studies are discussed, usually only the combined associations are discussed. In addition, the full range of individual-city associations should be documented. When seasonal results are discussed, usually summer or warm season and all-year results are presented. In addition, winter or cold season results whenever available should be documented. Specific examples are given below.

Revisions needed in the Figures The data presentation in the many figures in Chapter 6 provides a misleading impression of the overall patterns and consistency of the epidemiological data. By plotting only selected single-city associations, by omitting the

[28] 75 Federal Register 6501, February 9, 2010.
[29] 75 Federal Register 35548, June 22, 2010.
[30] 75 Federal Register 2997, January 19, 2010.

wide range of individual city associations in multi-city studies, by including meta-analyses that are known to overestimate positive associations due to publication bias, and by providing only limited information on the associations with other pollutants in the same or other studies, the ISA gives a false impression of the consistency of the data. Instead, the full range of individual-city associations in multi-city studies should be shown. Again, specific examples are given below.

Revisions needed in the integrative synthesis The current text is written with a bias to include all the arguments for ozone causing health effects. Instead, the text should weigh the evidence, both pro and con for ozone causing various effects at or below the current NAAQS. The current integrative synthesis ignores the issue of dose plausibility, leaves out consideration of personal exposures to ozone, and omits discussion of whether other pollutants might cause the same effect. As such it overstates the case for ozone having independent effects.

People are exposed to a wide variety of gases and particles in both indoor and outdoor environments. The lung has various defense mechanisms that help it continuously deal with these materials. The lung and respiratory system typically deals with a wide range of inhaled gases and particles without activating an inflammatory response. However, it is also capable of responding with an inflammatory response to a serious infective challenge. Ozone is both a natural constituent of the atmosphere and a pollutant. Although ozone is very toxic at high levels, it is also a natural constituent of the atmosphere with typical personal exposures below the outdoor background level. The issue for the ISA is at what doses does ozone overwhelm the body's defenses and cause concern. Therefore, a detailed discussion of the various potential effects in the context of the defense mechanisms and dose plausibility should be a major part of the integrative synthesis.

Revisions needed in Chapter 3 The discussion of ambient mixtures should include a discussion of seasonal differences including both other criteria pollutants and other photochemical oxidants. Although the ISA title refers to ozone and other photochemical oxidants, there is little or no discussion of the other photochemical oxidants in the draft. The section on ozone measurements indicates that other oxidants are not measured in routine monitoring networks, but that data are available from specialized studies. Because of the potential importance of the role of other photochemical oxidants in health effects, the ISA should include estimates of the concentrations of other oxidants from both measurements and photochemical model calculations. Estimates should be provided by season and location, since warm season associations are cited at many places in the ISA as being higher than all-year or cold season associations.

An additional important consideration is the level of ozone and other pollutants. In the U. S., peak ozone levels and the levels of other pollutants have been decreasing for over 40 years. Thus, the interpretation of older studies should include consideration of the higher concentrations to which the population was exposed at the time of the study. In order to provide such information for ozone, the long-term trends in ozone should be included in the ISA.

B. Comments on the data for specific endpoints

In the following, the epidemiological evidence for the most important heath endpoints is discussed and related to what is known about ozone effects from controlled studies. In the first sub-section, the evidence with regard to lung function and inflammatory markers is discussed. Next, there is a separate sub-section discussing the results of the multi-continent APHENA study that evaluated ozone associations with hospital admissions and mortality. The APHENA study is important because it is the largest multi-city study available and is discussed at several points in the ISA. This is followed by sub-sections discussing other endpoints discussed in the ISA.

Comments on respiratory effects – Section 6.2 of the ISA

Lung function effects The ISA presents the results for ozone/lung function associations but neglects to point out that many of the studies evaluated other pollutants and report many similar associations for those pollutants in single pollutant models. For example, the O'Connor et al., 2008 study evaluated five pollutants including ozone in a group of 861 asthmatic children in seven U. S. inner-city communities. The authors report stronger and significant positive associations of lung function parameters with three other pollutants compared to ozone in single-pollutant models. For asthma symptoms and missed school days, other pollutants also had stronger associations than ozone. Thus, the ISA gives a misleading impression of the role of ozone in the air pollution mix with regard to lung function and other respiratory effects.

In addition, the normal procedure of evaluating multiple lung function parameters at multiple lags and then reporting only the strongest associations increases the risk of false positives being highlighted in the ISA. For example, Pellegrino et al., 2005 warn that when too many indices of lung function are tracked simultaneously, the risk of false-positive indications of change increases.

Although there are many small positive associations of ozone with changes in lung function in the literature, the data are less consistent than indicated in the ISA. A particular important study was carried out by the Health Effects Institute in the Los Angles Basin, the area of the country with the highest ambient ozone concentrations. Avol et al., 1998 concluded that the relationships between ozone and pulmonary function were erratic and difficult to reconcile with existing knowledge about the acute respiratory effects of air pollution. In addition, the small changes in lung function that have been reported, to the extent they may be caused by ozone, are not medically significant given the transient, reversible nature of ozone lung function changes.

The ISA notes that newer data on children attending camps, outdoor workers, and other healthy populations were limited, and across these studies, ambient O_3 exposure was associated with both decreases and increases in lung function.[31] It goes on to note that a

[31] ISA, supra note 1, at page 6-17.

large number of older studies comprise a majority of the supporting evidence from epidemiology regarding lung function test effects, whereas recent studies, which were far fewer in number, provide less compelling evidence. Whether this is due to reduced ozone exposures, differences in study design, or other factors should be discussed in the ISA and considered in the integrative sections.

Inflammatory markers As with lung function measurements, studies of ozone association with the presence of inflammatory markers suffer from limitations due to the presence of other pollutants and multiple comparisons. The ISA also notes that the clinical relevance of most biomarker changes is not clear. The text in the ISA notes several additional reasons why there may be inconsistencies in the data. On balance, there is little evidence of significant associations with ozone in Figures 6-10 and 6-11. In addition, a number of these studies were conducted in Los Angeles and Mexico City where the subjects are exposed to high concentrations of both ozone and many other pollutants and report positive associations with various pollutants.

A particularly important study is described in the ISA as a well-designed panel study, Ferdinands et al., 2008. In this study, 16 adolescent long-distance runners in Atlanta, GA, were followed before and after exercise for 10 days in August 2004. Effect estimates for lags 0, 1, and 2 were positive, indicating O_3-associated decreases in airway inflammation. This study is important because the subjects, setting, and exercise level are just where one would expect to see ozone-induced inflammatory changes based on the clinical studies. Another study by Chimenti et al., 2009 measured some biological changes in adult male runners before and after races. However, the authors concluded that since no relationship was observed between neutrophil counts and inflammatory mediators 20 h after races, airways inflammation at this time point appears blunted in healthy runners and little affected by exposure to mild seasonal changes and airborne pollutants. Thus, under the situation with the greatest likelihood of inflammatory changes caused by ozone, there is little evidence of effects.

The lack of consistent increases in subclinical inflammatory markers is important information for the integrative synthesis. The lack of substantive effects in heavily exercising subjects suggests that there is even less likelihood of inflammatory changes due to ozone in the rest of the population as is goes about its daily activities. The findings in Adamkiewicz et al., 2004 of no inflammatory changes associated with ozone in elderly subjects including those with asthma and COPD confirm this view.

Comments on the APHENA study

The combined results of the large and comprehensive APHENA study are not consistent with ozone having a causal role in mortality or morbidity below the current standard. In October, 2009, the Health Effects Institute (HEI) published the results of the *Air Pollution and Health: A European and North American Approach (APHENA)*[32] study. The APHENA project was designed to take advantage of the largest databases available.

[32] Katsouyanni K. and Samet, J. (2009). Air Pollution and Health: A European and North American Approach (APHENA), HEI Report 142, October, 2009.

These had been developed by the three groups of investigators for earlier studies: 1) the *Air Pollution and Health: A European Approach* Phase 2 (APHEA2) study involving 32 cities; 2) the National Morbidity, Mortality, and Air Pollution Study (NMMAPS), conducted in the 90 largest U.S. cities; and 3) multicity research on the health effects of air pollution in 12 Canadian cities. Each database included air pollution monitoring data for particulate matter and ozone, health outcome data in the form of daily mortality for all ages, for persons younger than 75 years, and for persons 75 years or older (from all nonaccidental causes [all cause]), cardiovascular disease, or respiratory disease) and daily hospital admissions for persons 65 years or older (for cardiovascular and respiratory disease). Other database variables used for APHENA included weather data and a number of socioeconomic and other variables known or suspected to influence mortality or hospital admissions.

In the original studies, each of the three groups used different modeling methodologies and entered different variables into their models. Although each group found positive and significant relationships between PM_{10}/O_3 and mortality and some morbidity endpoints, the magnitude of the relationships differed by geographic region. One goal of APHENA was to use common methodologies and variables and reanalyze their data sets. They intended to create a central repository for all three of the time-series databases and use a common quality assurance approach. In addition, they would conduct analyses on a combined, pooled dataset to study a variety of sensitivity issues including effect modification. They would then investigate the sensitivity of the estimates to a variety of smoothing methods and to the number of degrees of freedom. They also intended to explore reasons for the geographical heterogeneity of the effect estimates seen in their original studies. Another important goal of the program was to understand the extent of coherence between mortality and hospitalizations using data from cities in North America and Europe.

In the original analyses, all three groups used a two-stage approach. In the first stage, risks were estimated for the individual cities, and in the second stage, evidence across the cities were combined. Each group used different methods to perform both stages in the original analyses. In APHENA, the investigators wanted to identify a preferred way to do both stages and apply common methodologies to the three data sets. For the first stage, they identified two smoothing techniques, natural splines (NS) and penalized splines (PS), and decided to use a number of degrees of freedom choices. They chose to use 3, 8 and 12 degrees of freedom and also the number of degrees of freedom chosen by minimizing the partial autocorrelation function (PACF).

For the second stage analyses, the two approaches used in original NMMAPS and the European studies represented the two major approaches used at the time to pool estimates. NMMAPS used Bayesian hierarchical regressions models while the Europeans used metaregression models. However, they could not determine which was the best method, so they decided to use the models interchangeably.

Using the two smoothing techniques together with the four choices for the degrees of freedom and three choices of lags (0-1 day, 1 day and distributive lags which provided

the cumulative effects of days 0 through 2) for each health outcome, the investigators ran a total of 24 different models for ozone. In addition, subsets of these choices were also used to examine the effects of controlling for PM_{10} and seasonal variations.
The results showed that the differences between the PS and the NS were very small in most cases and that the number of degrees of freedom tended to give similar results when greater than 6-8 degrees of freedom where used.

The overall modeling results for the mortality models and the morbidity models are summarized in Table 1 and 2, respectively. The denominator in the tables is the total number of different models that were run for each health effect outcome examined and the numerator is the number of models that resulted in a positive and statistically significant relationship between ozone and the health effect outcome. The way to interpret these tables is as follows. High ratios are suggestive of a robust and consistent relationship while low ratios are suggestive of no significant relationship. Intermediate values of the ratio suggest inconsistent and non-robust relationships that are dependent upon the model selected. Since there is no a priori way to determine the "correct" model, it is not possible to determine whether a small number (low ratio) significant and positive relationship represents real causal relationship or if they are false positives that can occur by chance or by confounding.

The all cause, all ages mortality results indicate a consistent relationship with ozone in Canada but somewhat less consistent relationships in Europe and the US. When the results for the two different age groups are examined, the interpretation of the results becomes even less clear. For $\geq$ 75 years of age, a consistent relationship still holds in Canada, but the European and US relationships become less consistent. When compared to the results for the < 75 years of age group, the results are implausible as they suggest that ozone is affecting the younger group more than the older group which goes against conventional wisdom. Controlling for PM makes the positive relationship for the older group disappear in all three locations, but the positive effect remains for the younger group except in the US where no relationship is evident. At all three locations a consistent summertime relationship is seen but vanishes in Europe and the US when PM is controlled. PM controlled model results were not presented for the Canadian data. In any event, the results are not consistent with the existence of a causal relationship between ozone and all cause mortality.

The cardiovascular mortality/ozone modeling results are somewhat confusing. A clear positive relationship was found only in Canada and only for the $\geq$ 75 years of age group. Few significantly positive relationships were found for either age group for the other locations and no relationship was found in Canada for the younger age group. When PM is controlled for, few significant relationships remain. The summer only results suggest significant relationships in Europe and the US, but they vanish when PM is controlled. Taken altogether, these results do not support a causal relationship between ozone and cardiovascular mortality when the models are controlled for PM.

The cardiovascular hospital admissions/ozone results are also confusing. The annual results show a few significant model-dependent relationships in Canada and the US but

none in Europe. When PM is controlled for, a few significant, model-dependent relationships remain in Canada, disappear in the US, but become consistently significant in Europe. The European results defy logic and were dismissed by the APHENA authors as a strong positive relationship was evident for respiratory hospital admissions and PM_{10}. The summer only results at all three locations show no significant relationships. Thus the weight of evidence from these results is consistent with the mortality results and does not suggest a causal relationship between ozone and cardiovascular hospital admissions.

In contrast to the cardiovascular mortality results, the respiratory mortality modeling results consistently show no relationship with one exception. None of the annual results at any location show any significant relationship between ozone and respiratory mortality. However for the summer, consistent significant results are found but only in Canada. Significant model-dependent results are seen in Europe and the US, but they disappear when controlled for PM. PM controlled results for Canada were not presented. Nevertheless, the weight of evidence of all the ozone/respiratory mortality model results does not support a causal relationship.

The respiratory hospital admissions show consistent significant relationships with ozone in Canada that disappears when PM is controlled. In the US and Europe, a few significant, model-dependent relationships are seen that persist when PM is controlled. However, during the summer when ozone is the highest and the strongest relationships would be expected, no significant relationships are found in either the US or in Europe. Consequently, the weight of evidence does not support a causal relationship between ozone and respiratory hospital admissions.

In summary, the APHENA results do not support EPA's claims of causal relationships between ozone and mortality or between ozone and hospital admissions.

Cause of Death	Canada	Europe	United States
All Cause – all ages	24/24	15/24	12/24
≥ 75 yrs	23/24	2/24	6/24
< 75 yrs	18/24	22/24	10/24
All Cause PM controlled – all ages	4/8	8/16	0/16
≥ 75 yrs	0/8	3/16	0/16
< 75 yrs	5/8	14/16	0/16
All Cause – summer only	9/9	18/18 (4/12)*	18/18(0/12)*
Cardiovascular – ≥ 75 yrs	24/24	3/24	2/24
< 75 yrs	0/24	8/24	2/24
Cardiovascular –PM controlled ≥ 75yrs	0/8	0/16	0/16
< 75 yrs	0/8	5/16	2/16
Cardiovascular – summer only	0/6	8/12(0/8)*	11/12(0/8)*
Respiratory – all ages	0/24	0/24	0/24
≥ 75 yrs	0/24	0/24	0/24
Respiratory – PM controlled – all ages	0/8	0/16	0/16
≥ 75 yrs	0/8	0/16	0/16
Respiratory – summer only	6/6	4/12(0/8)*	2/12(0/8)*

*Denotes the PM controlled ratio

Table 1: APHENA modeling results for mortality. The numerators represent the number of models that showed a positive and statistically significant relationship between O_3 and mortality while the denominator is the total number of models run.

Type of Admission	Canada	Europe	United States
Respiratory	18/24	8/24	7/23
Respiratory – PM controlled	0/8	7/16	5/16
Respiratory – summer only	3/3	0/4	0/4
Cardiovascular	5/24	0/24	3/24
Cardiovascular – PM controlled	3/8	16/16	0/16
Cardiovascular – summer only	0/4	0/4	0/4

Table 2: APHENA modeling results for hospital admission for patients 65 years and older. The numerators represent the number of models that showed a positive and statistically significant relationship between O_3 and admissions while the denominator is the total number of models run.

Comments on other respiratory endpoints

Respiratory hospital admissions In contrast to the summary of the APHENA study in the previous section, the ISA uses selected APHENA results to suggest generally positive associations with respiratory hospital admissions even though many of the associations shown are not statistically significant. The ISA also discusses several other multi-city studies. However, the fact that those studies reported associations of other pollutants with respiratory hospital admissions is not drawn out. For example, the Cakmak et al.,

2006 study of respiratory hospital admissions in 10 large Canadian cities reported positive associations for the four gaseous pollutants evaluated in single-pollutant models. Cakmak et al. evaluated associations for daily lags from 0 to 5 days and chose the lag with the strongest positive association for each city to include in the combined associations they report. Goodman[33] cautions that this can lead to bias. He notes that investigators tend to report, if not believe, the analysis that produces the strongest signal; and in each single-site analysis, there are various model choices that affect the estimated strength of that signal.

Others have also pointed out the critical importance of model choice, particularly when effect estimates are small. For example, Smith et al. caution:

> From a statistical point of view, the common epidemiological practice of choosing variables (including lagged variables, co-pollutants, etc.) that maximize the resulting effect estimates is a dangerous approach to model selection, particularly when the effect estimates are close to 0 (i.e., RR close to 1).[34]

Smith et al. note that Lumley and Sheppard, 2000 showed that the effect of choosing lags in this fashion has a bias which is of the same order of magnitude as the relative risk being estimated.[35]

The individual-city results were not reported by Cakmak et al., 2006. However, in studies that did report individual city results, the overall range among the cities was very wide. For example, the Medina-Ramon et al., 2006 study of 36 U. S. cities plotted the range of individual-city associations for the combined warm season ozone associations that they reported were statistically significant. The individual-city associations for COPD hospital admissions ranged from – 30 % to + 40 % for a 0.030 ppm increase in 8-hour ozone. The individual-city associations for pneumonia hospital admissions ranged from – 15 % to + 20 % for a 0.030 ppm increase in 8-hour ozone. When this wide range of individual-city results is considered in relation to Figure 6-19, it is clear that the current data presentation in Figure 6-19 is misleading. Thus, when all the individual-city results are considered, a very different pattern emerges making it very difficult to claim that there is a consistent ozone association with respiratory hospital admissions.

The combined associations for the two respiratory categories were positive in the Medina-Ramon study in the warm season, but were negative in the cold season and not significant over all the year. It is difficult to rationalize this pattern as an effect of ozone. It is not plausible that ozone would cause hospital admissions in one season and protect

---

[33] S. Goodman, *The Methodologic Ozone Effect*, Epidemiology, 16: 430-435 (2005).
[34] R. Smith, P. Guttorp, L. Sheppard, T. Lumley and N. Ishikawa, *Comments on the Criteria Document for Particulate Matter Air Pollution*, Northwest Research Center for Statistics and the Environment Technical Report Series No. 66, July 2001.
[35] T. Lumley and L. Sheppard, *Assessing seasonal confounding and model selection bias in air pollution epidemiology using positive and negative control analyses*, Environmetrics, 11, 705-717 (2000).

against hospital admissions in another season. It is not plausible that ozone would have a strong effect on hospital admissions in some cities and have a strong protective effect against hospital admission in other cities. Given the stochastic variability and the similar pattern of associations for many pollutants, it is clear that time series analyses even with massive databases is a blunt tool that does not allow one to ascribe effects to individual pollutants.

The ISA also notes that studies that focused on respiratory-related outpatient or physician visits found no evidence of an association with short-term O_3 exposure. This finding is also not consistent with ozone having an effect on hospital admissions, since if ozone were causing exacerbations of respiratory problems there would be more instances of out-patient and physician visits associated with ozone than instances of hospital admissions associated with ozone.

Respiratory mortality

The ISA in Section 6.2.8 notes that the data from the 2006 CD was inconsistent for an acute effect of ozone on respiratory mortality. The ISA goes on to indicate that the APHENA study found consistent positive associations for respiratory mortality in all-year analyses with stronger associations in analyses restricted to the summer. It also notes that respiratory mortality risk estimates were robust in the U.S. dataset in summer season analyses. In contrast, the HEI Review Committee's commentary on the APHENA study indicates that in all-year analyses the associations between ozone and respiratory mortality were generally close to zero and were not significant in any region or in the combined estimate for all three regions.[36] The APHENA investigators indicated that, generally, there was little evidence of an effect of ozone on respiratory mortality in any center.[37] While the associations were generally higher in summer-only analyses, only 2 of 12 model combinations were statistically significant and, when controlled for PM10, none of the 8 model combinations presented in the APHENA report were statistically significant. Thus, the ISA overstates the case for an effect of ozone on respiratory mortality.

Comments on cardiovascular effects – Section 6.3 of the ISA

The ISA notes that the cardiovascular morbidity data is inconsistent. Given the stochastic variability inherent in such studies, the many endpoints evaluated in even a single study, the complex mixtures involved, and the role of publication bias, it is not surprising that there would be some positive results for any endpoint. However, the overall pattern of ozone associations with cardiovascular morbidity endpoints is mixed and inconsistent. For example, the APHENA results for U. S. cardiovascular disease hospital admission associations with ozone (controlling for PM10) are essentially null. None of the 16 model combinations are statistically significant and 7 of the 16 model combinations actually have negative coefficients.

---

[36] APHENA report, supra note 32, at page 103.
[37] Ibid., at page 50.

The ISA claims that there is a consistent cardiovascular mortality signal, but the overall spatial and temporal mortality pattern, as shown below, is not consistent with causality. In addition, a mortality signal in the absence of a morbidity signal would be incoherent. Furthermore, the pattern of mortality associations for other criteria pollutants is remarkably similar to that for ozone in single-pollutant models.

The ISA also refers to toxicology studies showing ozone/cardiovascular effects. However, the animal studies cited used very high in vivo exposures where the respiratory defenses would be overwhelmed and in vitro exposures that the ISA acknowledges are speculative. In addition, the finding of increased aortic atherosclerotic lesion area in ApoE-/- mice reported in Chuang et al., 2009 is not unique to ozone. Similar findings have been reported in this sensitive animal model for elevated concentrations of both particles and other gaseous pollutants. Thus, the overall evidence for cardiovascular effects from current ambient ozone concentrations is weak and inconsistent.

Comments on all-cause mortality – Section 6.6 of the ISA

In addition to the discussion of respiratory and cardiovascular mortality in Sections 6.2.8 and 6.3.2.8, there is a major section on mortality that discusses all-cause mortality as well as specific causes of mortality. The ISA indicates that new multicontinent and multi-city studies reported consistent positive associations between short-term O_3 exposure and all-cause mortality in all-year analyses, with additional evidence for larger mortality risk estimates during the warm or summer months. It also indicates that these associations were reported across a range of ambient O_3 concentrations that were in some cases quite low. These statements overstate the case for consistent relationships. For example, the combined U. S. ozone associations for all-cause (and all-age) mortality in the APHENA multicontinent study are shown in Table 21 of the APHENA study. Of the 24 model combinations evaluated in single-pollutant models, 10 were not statistically significant and 7 were actually negative. More importantly, in the 16 model combinations presented that controlled for PM10, 4 were actually negative and none of the 16 were statistically significant. Thus, the APHENA ozone associations for all-cause mortality are not consistently positive.

The ISA supports the contention that there are consistent positive associations by referring to Figure 6-27. Figure 6-27 is misleading for several reasons. It includes only selected APHENA results. It includes three meta-analyses that are known to have substantial publication bias inflating the ozone/mortality association. It omits information on the full range of individual-city associations. Figure 6-29 of the ISA shows that the individual-city associations in the Franklin and Schwartz 2008 multi-city study range from -5 to +5 % for a 0.010 ppm ozone increase. That range translates into a - 10 % to + 10 % range for the scale of Figure 6-27, which is a 0.020 ppm increase in 24 –hour ozone. Similarly, the range in individual-city ozone associations in the Smith, Xu, and Switzer, 2009 study shown in Figure 6-28 and in the Bell and Dominici, 2008 study in Figure 6-31 translate into a range from roughly about – 10 % to + 10 % for the scale in Figure 6-27. Thus, if all the available data were shown on Figure 6-27, the perception of consistency would be dramatically altered.

The patterns in multi-city studies have spatial and temporal characteristics that are not consistent with causal relationships, demonstrating substantial stochastic variability. For example, the Smith, Xu, and Switzer[38] analysis of ozone mortality associations shows a wide range in individual-city associations as well as large regional differences in combined analyses that are not consistent with ozone causality. Figure 3 (Figure 1 from Smith Xu, and Switzer) is shown below. A similar pattern of wide variations from negative to positive in individual-city associations is evident in all the multi-city studies for which the individual-city associations are reported or plotted. The pattern in multipollutant analyses is not as well established, but there is substantial evidence of unrecognized stochastic variability in this case, too. The variability from city to city is too great to represent just heterogeneity due to different mixtures and exposures.

[38] R. L. Smith, B. Xu, and P. Switzer, Reassessing the Relationship Between Ozone and Short-term Mortality in U.S. Urban Communities, Inhalation Toxicology, 29(S2), 37-61, 2009.

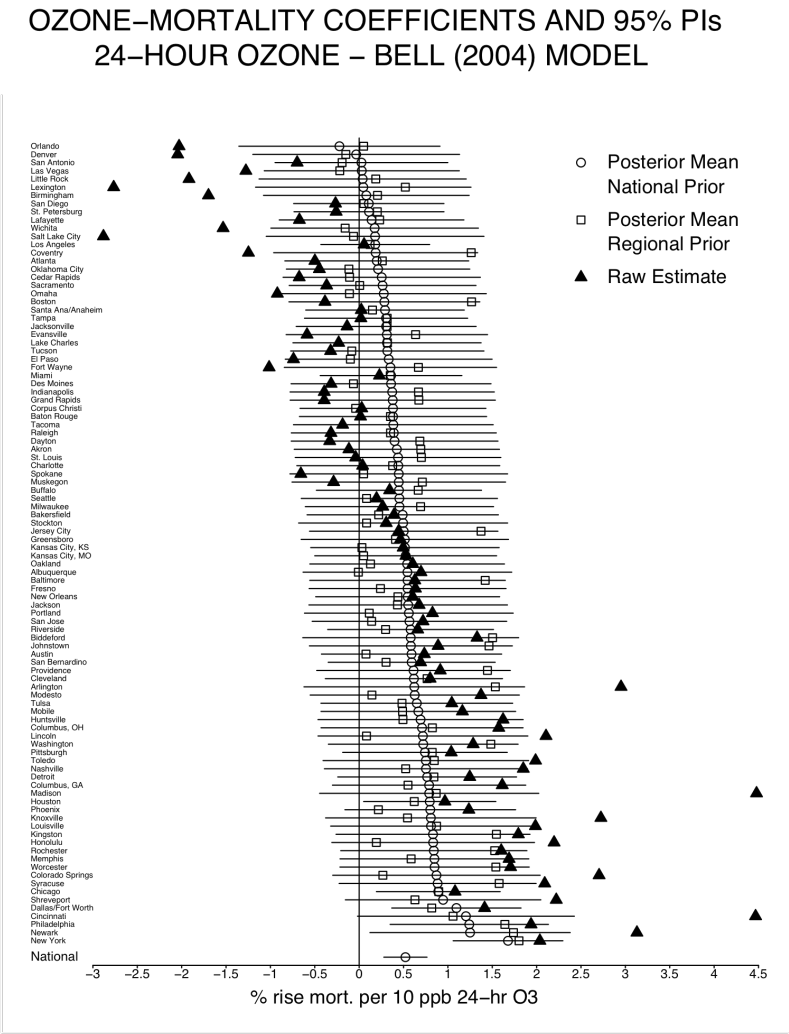


Fig. 1: 95% posterior intervals for the ozone-mortality coefficients, all-year data, by the hierarchical Bayesian method as in Fig. 2 of Bell et al. (2004). The Bayesian posterior estimates under the "national prior" (circles) are shown alongside those for the "regional prior" (squares) and the raw maximum likelihood estimates (triangles).

Figure 3 – Figure 1 from Smith, Xu, and Switzer, 2009

The spatial or regional patterns in ozone associations that are summarized in the ISA in Figures 6-32 and 6-33 and in Table 6-40 clearly are not consistent with the locations of highest man-made ozone in the monitoring data. The seasonal pattern with combined negative associations in colder months also raises a major issue with the claim of ozone causing mortality. The Schwartz, 2005 multi-city study noted the negative association in winter and suggested that it may reflect the negative association between wintertime ozone and primary air pollutants. Schwartz went on then to ask "might not the positive association in the summer likewise reflect confounding with some other pollutant?" These and other possible explanations need to be rigorously considered by EPA.

In discussing the shape of the concentration-response function, the ISA points out that combined mortality effects for ozone have been found at concentrations well below the current standard citing a multi-city study where high ozone days have been excluded, Bell et al., 2006. However, there is a follow-on study by Bell et al., 2007[39] that illuminates this issue. When Bell et al., 2007 restricted the analysis to days with low ozone, in order to see if the small combined association persisted, the range in individual-community associations widened. For example, when the data was restricted to days with ozone less than 0.02 ppm, the range in individual city mortality associations for a 0.01 ppm increase in ozone was from – 20 % to + 30 %. It is inconceivable that such low ozone exposures would be causing a dramatic increase in mortality in one city and protecting against mortality in another. With such wide variation, the interpretation of a small combined association as a health effect is highly questionable, especially in light of the fact that ozone indoors, where people spend about 90 % of their time is reduced about half or more by deposition to building surfaces.

In discussing the evidence for possible thresholds, the ISA downplays the findings of the Stylaniou and Nicholich, 2009 study noting that given the city-to-city variation in risk estimates, combining the city-specific estimates into an overall estimate complicates the interpretation of a threshold. By the same reasoning, given the city-to-city variation in risk estimates, combining the city-specific estimates into an overall estimate complicates the interpretation that the multi-city time series studies show ozone as having a causal effect on mortality.

A 2006 paper by Keatinge and Donaldson,[40] that is not included in the ISA, provides important new insights into the issue of modeling weather effects in ozone studies. The authors evaluated whether mortality that is often attributed to ozone and other pollutants in hot weather results from confounding by neglected weather factors. Their analysis was restricted to days when the mean daily air temperatures exceeded 18 degrees C in Greater London from 1991 to 2002, and evaluated mortality counts at an age greater or equal to 65. The adjustment for acclimatization was based on the characteristic pattern that has been reported by various investigators that the rise in mortality on hot days is followed by a prolonged reduction in mortality lasting at least 14 days. When only current temperature (average of days 0 to –2) was considered in the model, significant mortality was attributed to ozone. When they allowed for cumulative exposure to heat throughout the summer and for sunshine (which contributes to heat stress at any given temperature), the ozone association was reduced by a factor of 10 and was no longer statistically significant. This study indicates that previously neglected weather factors may be confounding the mortality analyses relied on in the ISA. It was noted in the 2006 Ozone CD that variations in treatment of weather can change the results by a factor of 2 and that

---

[39] M. L. Bell, J. Y. Kim, and F. Dominici, Potential confounding of particulate matter on the short-term association between ozone and mortality in multisite time-series studies, Environ Health Perspect, 115, 1591-1595 (2007).

[40] W. Keatinge and G. Donaldson, *Heat acclimatization and sunshine cause false indications of mortality due to ozone*, Environmental Research, 100, 387-393 (2006).

publication bias can inflate the perceived association by a factor of 3. The Keatinge and Donaldson analysis suggests that previously overlooked weather factors can reduce the association by a factor of 10.

With regard to uncertainty due to model selection, Koop and Tole, 2004[41] in another study not cited in the ISA conclude:

> Point estimates of the effect of numerous air pollutants all tend to be positive, albeit small. However, when model uncertainty is accounted for in the analysis, measures of uncertainty associated with these point estimates became very large. Indeed they became so large that the hypothesis that air pollution has no effect on mortality is not implausible. On the basis of these results, we recommend against the use of point estimates from time-series data to set regulatory standards for air pollution exposure.

The fact that the uncertainty due to model selection is much larger than the typical confidence limits on any given statistical association should be acknowledged in the ISA and considered in the interpretation of the epidemiological data. Given that the small positive results from time-series studies may reflect residual bias of the models due to weather, temporal or other unaccounted confounding factors, EPA cannot and should not draw conclusions on causality from these studies.

For all these reasons, the ozone mortality data is less consistent and persuasive than the ISA indicates and the conclusion that "there is likely to be a causal relationship between short-term O_3 exposure and all-cause mortality" cannot be supported.

Other endpoints

There are a variety of other acute endpoints discussed in Chapter 6 of the ISA. However, the data are cited as being suggestive, inconsistent, and/or inadequate. Because of the issues of stochastic variability, publication bias, model selection uncertainty, and potential confounding, it is not surprising that there would be some positive associations of ozone with any endpoint evaluated even in the absence of a causal relationship.

Comments on chronic effects – Chapter 7 of the ISA

Chapter 7 on chronic effects draws the conclusion that there is now substantial evidence for effects associated with chronic exposure. However, the evidence is less certain than indicated in the ISA. The data on lung function growth is mixed and inconsistent as indicated in the ISA, making it difficult to determine independent effects of ozone or other pollutants. What is described as the strongest new evidence is (1) one new chronic mortality study and (2) several studies implicating ozone in new onset asthma. All these studies are weaker than they appear because they rely on ambient ozone measurements to

[41] G. Koop and L. Tole, Measuring the Health Effects of Air Pollution: to What Extent Can We Really Say that People are Dying from Bad Air, J. of Environmental Economics and Management, 47, 30-54. (2004).

characterize the exposures. In reality, personal exposure is the proximate cause of any chronic effects and total personal exposure is the metric that should be considered.

For new onset asthma, the studies involve analyses of data from the Southern California Children's Health Study. Studies of effects in Los Angeles or Mexico City, areas with high historic levels of ozone and other pollutants, are not relevant to effects at or below the current standard. Another major concern with theses studies is that the many other variables that are known to affect asthma are not fully considered. For example, differences in living conditions in the indoor environment (or changes in living conditions in the indoor environment due to energy conservation) may play a role by increasing exposure to common household allergens such as mold, dust mites and animal dander.

With regard to asthma and chronic ozone, Dockery et al., 1996[42] evaluated asthma prevalence and respiratory symptoms in a group of 13,369 8 to 12 year old children in 24 communities and reported that the prevalence of asthma or asthmatic symptoms was not associated with chronic exposure to ozone. The Dockery et al., 1996 study is not included in the ISA. In addition, McConnell et al., 2002 followed 3,535 children with no history of asthma from 12 communities in Southern California, as part of the Children's Health study. The risk for developing asthma was not greater in high ozone communities compared to low ozone communities. Although there has been an increase in asthma in both developed and developing nations in the past 20 or 30 years, outdoor air pollution, in general (and ozone, in particular), has been decreasing. Outdoor air pollution cannot be the cause of the increase in asthma.

With regard to chronic mortality, the ISA summarizes several studies showing no effect but focuses on the Jerrett et al., 2009 study as showing a chronic respiratory mortality signal for ozone. However, the respiratory mortality signal is present only for females in spite of the fact that males would be expected to receive higher ozone doses by being outside exercising more than females. In addition, the regional results reported by Jerrett et al. show no respiratory mortality effect in Southern California, the Northeast, or the Industrial Midwest, the regions of the country with the highest historic man-made ozone exposures. Finally, the presence of a chronic respiratory mortality signal is not coherent with the lack of an acute respiratory mortality signal in APHENA. For these reasons, the evidence for a chronic ozone mortality effect is much weaker than indicated in the ISA.

Although there are toxicological studies showing effects from chronic ozone exposures, these studies utilize high ozone exposures compared to the day-to-day personal ozone exposures of the population. At high chronic exposures, effects due to repeated lung injury and repair cycles are apparent. However, the relevance of these studies to the personal exposures of the population which are typically one-quarter of the ozone measured at ambient monitors is not clear.

---

[42] D. Dockery, J. Cunningham, A. Damokosh, L. Neas, J. Spengler, P. Koutrakis, J. Ware, M. Raizenne and F. Speizer, *Health Effects of Acid Aerosols on North American Children: Respiratory Symptoms*, Environmental Health Perspectives, 104: 500-505, 1996.

C. Overall comments on conclusions regarding causality

The ISA overstates the consistency and coherence of the available data with regard to respiratory morbidity and mortality and all-cause mortality. There are now numerous multi-city studies that all exhibit a wide range of individual-city associations between ozone measured a central city monitors and daily health statistic for mortality and morbidity. The overall patterns in the associations are not consistent with ozone causality. The individual city associations vary widely from positive to negative, a finding that is not addressed sufficiently in the ISA. It is not plausible that ozone would cause effects in some cities and protect from effects in other cities.

The seasonal variability in associations is also not consistent with causality. Combined associations in multi-city studies are generally negative in winter and positive in summer. The differences in ozone exposure between summer and winter are not that large as shown in the many graphs in Section 3.8.5 that they can explain the negative combined associations in winter. The correlations of ozone with other pollutants are very different between winter and summer as shown in Figure 3-44. This suggests that ozone is acting as an indicator of the air pollution mix in summer but is anti-correlated with the air pollution mixture in winter. The strong possibility that ozone is only showing associations because it is an indicator and not an independent causal factor needs to be acknowledged and fully discussed in the ISA.

The spatial variability of combined ozone associations in acute multi-city studies is also not consistent with ozone causality. Studies by Bell et al. and Smith et al. show that there are positive combined ozone associations in some regions but not in others. In fact, the spatial pattern of positive ozone associations is similar to that for particulate matter (PM) and is not consistent with the locations of highest man-made ozone exposure. These facts need to be acknowledged and discussed in the ISA.

Finally, the multicontinent APHENA study demonstrates an inconsistent and incoherent pattern of combined mortality and hospital admission associations. The HEI Review Committee's commentary, for example, indicated that it is remarkable how little coherence there is for the O_3 effects.[43] Even with a massive database, APHENA modeling shows few statistically significant combined effects. Although the APHENA study did not report all the individual-city associations, the small subset that is included in the report shows, in agreement with other multi-city studies, that the combined associations result from a mix of individual-city associations that include cities with strong positive ozone associations, cities with strong negative associations, and cities with no ozone association. The overall pattern in APHENA is not consistent with ozone causing respiratory or cardiovascular mortality and morbidity. Rather the pattern of combined associations probably represents residual confounding.

With regard to chronic mortality, the ISA focuses on the Jerrett et al., 2009 study as showing a chronic respiratory mortality signal for ozone. However, the respiratory

[43] APHENA report, supra footnote 32, at page 109.

mortality signal is present only for females in spite of the fact that males would be expected to receive higher ozone doses by being outside exercising more than females. In addition, the regional results reported by Jerrett et al. show no respiratory mortality effect in Southern California, the Northeast, or the Industrial Midwest, the regions of the country with the highest historic man-made ozone exposures. Finally, the presence of a chronic respiratory mortality signal is not coherent with the lack of an acute respiratory mortality signal in APHENA. For these reasons, the evidence for a chronic ozone mortality effect is much weaker than indicated in the ISA.