Review and Critique of U.S. EPA's Assessment of the Health Effects of Particulate Matter (PM)

by

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

Prepared for
The American Coalition for Clean Coal Electricity

August, 28, 2012

EXECUTIVE SUMMARY

Background

The 1970 Clean Air Act Amendments established the concept of National Ambient Air Quality Standards (NAAQS) and required that they be set for individual "criteria" air pollutants, which are those pollutants that are ubiquitous and are emitted from numerous or diverse mobile or stationary sources. Health-based primary standards and welfare- or ecology-based secondary standards had to be promulgated and the scientific basis for the standards has to be revisited and reevaluated every five years.

EPA has designated six criteria air pollutants: carbon monoxide (CO), nitrogen dioxide (NO_2), sulfur dioxide (SO_2), ozone (O_3), lead (Pb), and particulate matter (PM). The first five pollutants are specific chemical constituents while PM is a mixture of hundreds of different chemical species with differing degrees of toxicity and health effects. PM consists of a mixture of solid or liquid substances suspended in the air. It includes dust, smoke and condensed material. The terms particulate matter, particles and aerosols are sometimes used interchangeably.

Over the years the regulatory focus has shifted to different kinds of PM. Originally the focus was on total suspended particulates (TSP) defined as anything that was collected on a fiber glass filter in a High Volume Air Sampler. On average, these samplers would collect any particles in the air that were less than 30 - 40 micrometers (μm) in diameter although this would vary with both wind speed and direction. However, from a health perspective, only particles that were small enough to be inhaled were of concern. Consequently, in 1987 EPA promulgated the first inhalable PM NAAQS, which they called $PM_{10,}$ and it is composed of particles that have diameters of 10 μm or less (or about 1/7 the width of a human hair). In the 1990s, additional attention became focused on fine particles which could be inhaled deeply into the lungs. Consequently, in 1996, EPA promulgated the first fine particle or $PM_{2.5}$ NAAQS. $PM_{2.5}$ are particles that have a diameter of 2.5 μm or less (or about 1/30 the width of a human hair). The particles with a diameter larger than 2.5 μm but less than 10 μm are $PM_{10\text{-}2.5}$, which EPA has named the thoracic coarse fraction. In addition, there is another size class of particles called ultrafine particles (UFP), which are a sub-class of $PM_{2.5}$ particles with diameters less than or equal to 0.1 μm. At present, EPA has primary and secondary NAAQS for PM_{10} and $PM_{2.5}$. To determine compliance with a NAAQS, samplers have been deployed to collect the PM in these two size-fractions on filter media which are then weighed. Daily (24-hour) composite samples are used to determine compliance with the 24-hour NAAQS. The annual averages of the daily samples are computed to determine compliance with the annual NAAQS.

In considering the establishment of NAAQS, EPA relies on three types of health effect studies: controlled human exposures ("clinical"), animal toxicology ("toxicology") and epidemiology studies. In all NAAQS reviews prior to the 1996 PM review, EPA relied most heavily on controlled human exposures, which establish health effect endpoints as a function of exposure and demonstrate causality, and the toxicology studies which provide insights as to the mode of the damage caused by an exposure. Epidemiology studies were used if they supported the findings in the other two types of studies because epidemiology studies can only identify statistical associations between air pollutant concentrations and health endpoint incidence and cannot be used to demonstrate causality (cause-effect relationships). In the past, the discovery of

a moderate association between a chemical substance (whether a pollutant or a drug) and a health endpoint simply meant that additional investigations were warranted, such as clinical and toxicological studies.

For the PM NAAQS review that ended in 1996, EPA subordinated its reliance on human exposure and toxicological studies because they showed no evidence of effects at concentrations near the level of the existing NAAQS. Instead, they relied primarily on epidemiology studies, which were finding very weak statistical associations between measures of PM, including PM_{10} and $PM_{2.5}$, and mortality (death) at ambient concentrations well below the then existing PM_{10} NAAQS. EPA recognized that there were large uncertainties associated with the epidemiology studies because they cannot demonstrate cause and effect. Despite this realization, EPA promulgated new annual and 24-hour $PM_{2.5}$ NAAQS based on the epidemiology findings.

Between the 1996 review and the 2006 review, EPA funded ten years of research targeted at reducing these uncertainties but the research failed to do so. Instead, the intervening years produced numerous new epidemiology studies that continued to find associations between health effects including mortality and $PM_{2.5}$ and PM_{10}, as well as with the other gaseous criteria air pollutants. As a result in 2006, EPA used the epidemiology results to lower the 24-hour $PM_{2.5}$ NAAQS, and again, acknowledged that there were serious uncertainties with the epidemiology studies.

For the most recent NAAQS reviews for ozone, nitrogen dioxide and sulfur dioxide, EPA, for the first time fully embraced the epidemiology studies that implicated these pollutants to mortality and/or morbidity (illness). As a result, EPA promulgated a new, very stringent eight-hour NAAQS for ozone in 2008, and new, very stringent one-hour NAAQS for nitrogen dioxide and sulfur dioxide in 2010.

When the current ongoing PM review began in 2009, EPA largely ignored the uncertainties that still remained despite more than a decade of targeted research and concluded that the epidemiological evidence is sufficient to conclude that causal relationships exist between both acute (short-term) and chronic (long-term) exposures to $PM_{2.5}$ and mortality and cardiovascular and respiratory effects. As a result, they have proposed to lower the annual $PM_{2.5}$ standard even further.

Epidemiology Results

EPA's Interpretation

EPA's main conclusion is summarized in the preceding paragraph - that, despite uncertainties and the lack of clinical and toxicological evidence, epidemiology studies are sufficient to conclude that $PM_{2.5}$ causes adverse health effects at current concentrations in the ambient air. EPA bases this on "important new information" from "hundreds of new epidemiological studies conducted in many countries around the world." EPA claims that the scientific evidence "have undergone intensive scrutiny through multiple layers of peer review and opportunities for public review and comment." They state that: "$PM_{2.5}$ risk estimates were found to be consistently positive, and slightly larger than those reported for PM_{10} for all-cause, and respiratory- and

cardiovascular-related mortality." Thus, EPA presents a coherent and impressive argument to support their conclusions of causal associations between ambient $PM_{2.5}$ concentrations and mortality in both short- and long-term studies.

A Different Interpretation

Another interpretation of the scientific evidence leads to the conclusion that the results cited by EPA have been cherry-picked and conflicting evidence has either been ignored or misrepresented. In addition, the results fail to pass numerous reality checks. The reality checks include inconsistency between epidemiology and real-world dosimetry and toxicological evidence. Furthermore, the epidemiological studies EPA cites are inconsistent both across geographic areas and among socioeconomic groups, raising questions about EPA's reliance on them. These issues are discussed below.

Dosimetry

Information on the dosimetry of particles, that is the deposition, clearance, and retention of particles within the respiratory tract, is critical to understanding the health effects of inhaled particles because the cause of a biological response to PM is due to the dose deposited at the internal target site, rather than the external exposure. Information from dosimetry can answer key questions. The most important question is whether the doses of fine particles to target tissue in a 24-hour period or over a lifetime are high enough to cause the effects implied by the epidemiological associations.

There are several studies that address this. One study estimated that the mass of particles deposited per unit of alveolar–interstitial (deep lung) tissue in humans inhaling particle concentrations as high as 50 $\mu g/m^3$ for 24-hours (the 24-hour $PM_{2.5}$ NAAQS is 35 $\mu g/m^3$) was only in fractions of nanograms (10^{-9} gram) of particles per square centimeter. For individual particle components that are of interest as potential causal agents, much lower deposition levels were found. For example, sulfate deposits were only in the range of picograms (10^{-12} gram) per square cm of alveolar surface, and levels of elemental carbon, iron or trace elements were not higher than a fraction of a picogram per square cm of surface. For toxic metals, suggested as a probable cause of fine particle toxicity, the estimated 24-hour deposition levels were extremely low, not exceeding tens of femtograms (10^{-15} gram) per square cm of alveolar-interstitial surface. It is inconceivable that such small amounts of these materials could cause the effects implied by the statistical associations because these dosages are orders of magnitude lower than those that produced biological responses in toxicological studies. It is a challenge for toxicology to explain how such low doses of particles can be causing the health effects implied by the epidemiological associations.

Toxicology

Toxicology is known as the science of poisons, where a poison may be any substance which when acting directly through its inherent chemical properties is capable of destroying or seriously endangering life. Any substance, even food and water, may be harmful if absorbed or ingested in excessive amounts. The dose determines whether or not injury will occur, requiring the toxicologist to pay careful attention to the quantitative measurement of both dosage and

effect, before the delivered dose is declared as "harmful." One of the major uncertainties EPA acknowledged during the 1996 review of PM NAAQS was the lack of demonstrated mechanisms that would explain the mortality and morbidity effects implied by the epidemiological associations. A review of the toxicology material EPA used for the 2009 review reveals that, despite over a decade of expanded and focused research, there are still no data from controlled studies that indicate how anthropogenic PM at current ambient levels is causing the mortality and morbidity effects implied by the epidemiological associations that EPA relies on.

To evaluate the risks posed by Superfund sites, which can contain numerous hazardous chemicals, or to assess the hazard due to an exposure to a mixture of hazardous air pollutants, EPA would typically conduct a standard U.S. EPA health-risk assessment based on the relative toxicity of the mixture. Although EPA has not done this for PM, it would provide another reality check for the epidemiology results. However, such an assessment has appeared in the scientific literature (Valberg, 2004). In that assessment, the author used the chemical-specific, dose-response data typically used in U.S. EPA human health-risk assessments to evaluate the risk associated with a mixture of 27 separate chemical constituents typical of ambient PM with a total $PM_{2.5}$ concentration equal to that of the current annual $PM_{2.5}$ NAAQS of 15 $\mu g/m^3$. The assessment relied on established, no-effect thresholds for noncancer health endpoints. The author found that the chemicals identified as constituents of ambient PM are present at concentrations considerably below the regulatory thresholds (for which no adverse health effects are anticipated for a lifetime of exposure) used in risk assessment. From the perspective of risk assessment, the author concluded that, using EPA's own risk assessment methodology, exposure to the concentrations of chemicals that constitute ambient $PM_{2.5}$ (e.g., sulfate, nitrate, and 25 other constituents) cannot be expected to cause death. Hence, the author noted that the health effects attributed to ambient PM in the NAAQS review appear to be at odds with what would be predicted from a standard U.S. EPA health-risk assessment for PM chemicals. The author discusses several possible explanations for this paradox, including the implausible possibility that the toxicity of ambient PM is unrelated to its chemical constituents, or that PM mass concentration is not the causal factor in the reported associations. The current EPA rulemaking materials are silent on the existence of this paradox, much less on the possible explanation.

Another paradox is found in controlled toxicological studies that expose either animals or human volunteers to concentrated mixtures of ambient PM or aged power plant emissions. These studies, using much higher concentrations than are found in the ambient atmosphere, find no severe effects.

Despite the outpouring of toxicological studies evaluating many possible mechanisms by which $PM_{2.5}$ may cause harm, there is a lack of consistent findings of clinically-relevant PM effects at high concentrations. Thus, toxicology cannot explain how low concentrations can be causing death.

Epidemiology

There have been historical air pollution episodes like the London and Donora, Pennsylvania episodes where relationships between air pollution and acute health effects have been documented. In the U.S. today, levels of PM, even in cities with the highest concentrations are a fraction of what they were during these episodes. As a result, there is doubt in the minds of many air pollution professionals that today's levels of PM (PM_{10} or $PM_{2.5}$) in the U.S. are

causing adverse health effects.

A close examination of the results of the epidemiology studies cited by EPA reveals that they do not produce consistent results. EPA relies on two types of epidemiology studies: "long-term," or "chronic," studies in which health effects in cities with different levels of air pollution are compared over long periods of times; and "time-series," "short-term" or "acute" studies in which health effects (e.g., deaths, emergency room visits) are compared within a city as air pollution levels fluctuate. Many of the recent short-term studies have also been "multi-city" studies as they compare the single-city time-series results among cities and then pool these results to estimate regional or national averages.

The results for both types of "multi-city" studies indicate that:

- First, the results exhibit significant heterogeneity (i.e., they are inconsistent). For PM, the ranges of risks are implausible and inconsistent with a causal PM_{10}/mortality or $PM_{2.5}$/mortality relationship. The risks range from negative (a beneficial effect) in some cities to positive (a harmful effect) in other cities and most risks are not statistically significant. In fact, in one major study of the 90 largest cities in the U.S., positive, statistically significant correlations between PM and mortality were found in only two of the 90 cities.

- Second, the distribution of risks across all the cities is nearly the same for all the criteria pollutants - more than half exhibit a positive risk and a quarter to nearly half show a zero or negative risk. It seems irrational to single out PM as the causal agent when the results for the other criteria pollutants are nearly identical.

- Third, there is significant spatial heterogeneity in the associations, with no effect seen in western U. S. cities in most studies.

- Fourth, the PM risks typically disappear or become statistically insignificant when other criteria pollutants are included in the statistical models.

Taken together, these results do not create a picture of consistent causal relationships between PM and mortality or morbidity.

Why the Epidemiology Results Are Inconsistent

Publication Bias

Publication bias is a major issue in assessing the epidemiological literature. Publication bias occurs because authors are inclined to selectively pick the modeling results that show the largest effects and editors are more likely to publish papers with positive findings. Consequently, there will be more papers in the literature that show positive epidemiology results than those that show negative results. This has been identified as a major issue in air pollution epidemiology and has led to inflating the size of any potential effect.

Confounding

A confounder is an extraneous variable that correlates with both the dependent and independent variable. Such a relationship is termed a spurious relationship. In the case of a risk assessment, it is important to control for confounding to isolate the risk of a particular hazard. All epidemiology studies must deal with the issue of confounding. The ambient air can contain trace amounts of hundreds of chemical species both in the gas and particulate phase. Most epidemiology studies only focus on PM mass, but PM contains measureable amounts of nearly every element that exists in the earth's crust. Individual elements can exist as different chemical compounds. Consequently, there are hundreds of potential confounders in the air and only a tiny fraction of them are even measured. Because of this, in a study of any one component of air pollution, other components that may be associated with health impacts must be controlled. Very few studies do this for even the ones that are measured. This means that the potential for confounding by other substances in the atmosphere can never be completely controlled and their effects ruled out.

In all air pollution epidemiology, weather is also an obvious confounder. In addition, other temporal effects such as season, cyclic diseases, and day-of-the-week patterns must be controlled for. A prestigious Special Panel of the Health Effects Institute concluded however, that it is not known how to do this. They stated: "Neither the appropriate degree of control for time in these time-series analyses, nor the appropriate specification of the effects of weather, has been determined."

Many short-term, time-series studies of air pollution report associations between fine PM and various measures of human health, such as the number of daily deaths and hospital admissions, in single pollutant analyses. However, when possible confounding by other pollutants is explicitly addressed, many of the studies find no association between PM and measures of human health.

A study by Janes et al. (2007) concluded that the reported associations between PM and mortality may actually be attributable to inadequately controlled confounding. They hypothesized that the association between national trends in fine PM and mortality "is likely to be confounded by slowly time-varying factors, such as changes in industrial activities and the economy, improving health care, and large scale weather events." When these factors were removed, there was no evidence of an association between 12-month exposure to $PM_{2.5}$ and mortality.

While the time-series studies are based on day-to-day differences in ambient pollution levels, the long-term studies use differences in pollutant levels between cities. Many more confounders must be controlled in long-term studies than in time-series studies. For example, smoking, which is not a confounder in time-series studies because an individual's status as a smoker or non-smoker does not vary day-to-day, is an important potential confounder in long-term studies. Because air pollution is measured on a city-wide level, any factor that is associated with mortality and varies from city to city, such as life-style or socio-economic factors, is a confounder in the long-term studies. Control of confounding by such factors can be extremely difficult. Adjustment for socio-economic factors is particularly difficult, and there is little

assurance that residual confounding bias can be eliminated. For example, one prominent study finds an association between PM levels and mortality only for those with a high-school education or less. Without a known physiological basis for such an effect, this implies that control for some unknown lifestyle-related confounder is not adequate. Thus all long-term studies need to be reviewed with particular caution.

Failure to identify and control for confounders can lead to unrealistic results. In the Six Cities study, the $PM_{2.5}$ relative risk (RR) for all-cause mortality of 1.26 reported for residents of Steubenville, OH as compared to Portage, WI is implausibly large. Using the risks associated with smoking estimated in the Six Cities study shows that this RR is equivalent to increasing cigarette consumption among smokers by 25 pack-years (1 pack per day for 25 years). This is simply not biologically credible. Furthermore, the Six Cities study reports a RR of lung cancer mortality of 1.37 for residents of Steubenville as compared to Portage, which would make ambient $PM_{2.5}$ much more potent than direct emissions from coke ovens. This is also not biologically credible. (Moolgavkar, 2005)

Exposure Uncertainty

An ideal epidemiology study would have information on both exposure and disease outcome on each individual in the study. Most air pollution epidemiology studies do not even come close to having this type of information. Concentrations of pollutants in the air are obtained from air monitoring sites at locations scattered throughout the United States. These monitors have generally been sited to determine compliance with the NAAQS. In epidemiology studies, the concentrations measured at these sampling stations are assumed to be representative of actual exposures received by individuals living near them. This assumption is false. Pollutant concentrations can and do exhibit significant spatial (horizontal as well as vertical) and temporal gradients. Actual exposure is determined by where people live, where they work, the time they spend indoors and outdoors, and the myriad habits of daily life. In the 2006 PM review, EPA pointed out that people spend about 90% of their time indoors and only about 6% of their time outdoors where they are directly exposed to ambient PM. The general population spends between 50 and 60% of the time at their place of residence. For the frail population in hospitals and nursing homes, the time spent indoors approaches 100%. Thus, the exposure of individuals varies greatly, and central monitors do not capture this variation, creating significant uncertainty in exposure estimates in epidemiological studies.

Because of the way exposures are estimated in the short-term, time-series studies, there is no assurance that any of the individuals who died or became ill were actually exposed to the highest levels of pollution. In the long-term studies, there is no way of assessing the actual exposures of any individuals who died over the course of the study.

Further, recent studies have shown that exposure uncertainty or measurement error can produce false linear concentration-response functions that have no thresholds. Thus, epidemiology studies cannot inform us as to whether there is or is not a biologic gradient for ambient PM at low concentrations, or whether or not there is or is not a threshold. Consequently, EPA's claims of a no-threshold response down to zero PM are not substantiated.

Model Selection Bias

In epidemiology, statistical models are used to relate a health outcome to various factors that may contribute to the occurrence of that health outcome. Selecting an appropriate statistical model for epidemiological analyses of air pollution data is an extremely important process that can affect the outcome of the study in a very significant way. It can make the difference between finding a positive association, a negative association or no association. It involves making a number of choices which include:

- How is confounding by weather to be controlled? That is, what functional form should be assumed for the effects of weather variables, such as temperature and relative humidity?
- What weather variables should be used?
- What co-pollutants should be included and what averaging time should be used?
- What temporal effects need to be controlled and to what degree?
- What lag structure should be assumed? That is, how many days after exposure to a pollutant should one expect to see an effect on health?

There is little biological knowledge to inform these choices that must be made. Unfortunately, most investigators do not make these choices systematically and many choose the model that maximizes the effect estimates. Because of the large number of possible models, the results that are reported could have occurred by chance.

Researchers that have examined this issue in depth conclude that when the uncertainties introduced by model selection are considered, the uncertainties "become so large as to question the plausibility of the previously measured links between air pollution and mortality" (Koop and Tole, 2004). Others have concluded that even if the true effect of pollution is zero, the estimated effect may be positive because it is impossible to control temporal trends or weather without accurate information from external sources that does not exist.

An important paper on model selection bias that deserves attention is Koop et al. (2010). In this study, the authors conduct a comprehensive analysis of air pollution-morbidity relationships for eleven Canadian cities over a long record from 1974 to 1994. As a result, they have a unique data set that allowed the examination of both spatial and temporal variations. In addition to including the five criteria pollutants, CO, PM, SO_2, NO_2 and O_3, they also controlled for socioeconomic factors, smoking and meteorology. Much shorter subsets of this data set had been previously analyzed without the socioeconomic and smoking variables by a number of research groups to demonstrate significant relationships with a number of health outcomes and individual pollutants. The long data set enabled the present investigators to explore the impact of significantly lower air pollution concentrations at the end of the data set compared to the beginning. Koop et al. also employed the two major methods used to formulate the statistical models in time-series studies, model selection by the use of some statistical criteria and Bayesian Model Averaging (BMA), to address the all-important issue of model selection uncertainty.

The results of the BMA analyses show that the health outcomes are explained by the smoking and the socioeconomic variables and that none of the air pollutants showed a statistically positive relationship with health. This study demonstrates the importance of: 1) incorporating smoking

and socioeconomic variables into the models, 2) using a longer time series that has significantly different pollutant concentrations at the beginning and end of the study, 3) using the BMA approach which minimizes model selection uncertainties and finds insignificant health impacts.

Concluding Remarks

Because of the voluminous number of air pollution epidemiology studies that have been published, EPA is able to cherry-pick the results to develop a coherent and impressive argument to support their conclusions of causal associations between ambient $PM_{2.5}$ concentrations and both short- and long-term mortality. However, close scrutiny of these studies reveals inconsistent results that do not support a causal relationship between PM and the mortality/morbidity outcomes. In addition, despite 15 years of research targeted to demonstrate biological plausibility of the statistical associations, no mechanisms have been identified. Toxicology and clinical studies are inconsistent with a causal relationship at concentrations anywhere near ambient levels of PM in the U.S. A combination of model selection bias, confounding, exposure uncertainty and publication bias provides a more likely explanation for the observed statistical relationships. Taken together, we conclude that the epidemiological evidence relied on by EPA is scientifically unsound and should not be used as a reason to drive the NAAQS lower and lower.

I. Purpose and Scope of the Report

The purpose of this report is to provide sufficient background to understand the basics of the Particulate Matter (PM) issue including why EPA thinks that PM poses a serious health risk and why, in the opinion of other scientists, EPA has misinterpreted or misrepresented the scientific evidence. The report is broadly divided into three main parts. The first part provides a historical perspective on the health review process specified by the Clean Air Act and the evolution of PM health science. The second part examines the results of the health effect studies from two perspectives - EPA's and the scientists who are critical of EPA. Then the reasons for the differing opinions will be addressed. The views expressed in this paper do not necessarily represent the views of ACCCE or any of its individual members.

II. Historical Perspective

A. The Process Used by EPA to Set the National Ambient Air Quality Standards (NAAQS)

The 1970 Clean Air Act Amendments (CAAA) established the concept of NAAQS and required that they be set for the individual "criteria" air pollutants, which were those pollutants that were ubiquitous and were emitted from numerous or diverse mobile or stationary sources. Health-based primary standards and welfare- or ecology-based secondary standards had to be promulgated and the scientific basis for the standards had to be revisited and reevaluated every five years.

The CAAA also specify that the criteria developed in this five-year cycle and EPA's analysis of it be reviewed by the Clean Air Scientific Advisory Committee (CASAC),[1] who reports their findings and recommendations directly to the EPA Administrator. Since the process used in the most recent review cycles has changed somewhat, both will be briefly described. In the reviews of the PM standards completed in 1996 and 2006, EPA first would produce a Criteria Document (CD), which was a compendium of all the latest available science on PM. In the 1996 review, EPA had to produce multiple drafts of the CD before CASAC was finally satisfied that it contained and evaluated all the recent science. About the time that EPA produced the second draft of the CD, they also submitted to EPA their first draft of a Staff Paper (SP). The SP contained EPA's risk assessment of PM and EPA's preliminary recommendations on the form, averaging-time and range of levels for the PM NAAQS, which they justified based on the science in the CD. The SP also required several iterations before CASAC came to "closure" on the document. Once CASAC came to closure on the SP, EPA subsequently published a Notice of Proposed Rulemaking in the Federal Register (NPRM). Following a public comment period, EPA published the Final Rule. The 2006 review followed the same process.

For the current review, which is supposed to end in December 2012, the CD was replaced with a document called an Integrated Science Assessment (ISA) and the SP was replaced with two

[1] CASAC is a 7-member committee of independent air quality, health, and ecological effects scientists who are appointed by the EPA administrator to provide scientific advice to the Administrator. For each individual NAAQS review, between 10 and 15 additional scientists with expertise on that pollutant are added to form a CASAC Panel that conducts the NAAQS review.

documents, a Risk and Exposure Assessment (REA) and a Policy Assessment (PA). The ISA is supposed to be a more concise summary and evaluation of the science published since the previous review. The REA contains the risk assessment while the PA contains EPA staff's recommendation for the proposed standard and the scientific rationale. By design, EPA now limits CASAC's review to two drafts of the ISA, REA and PA. Although CASAC completed its final review of these documents on September 10, 2010, EPA did not issue a final ISA until April 2011. After they were ordered on May 31, 2012 by the DC District Court in response to a lawsuit initiated by the American Lung Association and others, did EPA issue the NPRM which states that they intend to issue a Final Rule in December 2012.

B. What is PM?

Particulate matter (PM) is a mixture of solid or liquid substances suspended in the air. It includes dust, smoke and condensed material. The terms particulate matter, particles and aerosols are sometimes used interchangeably. Over the years the regulatory focus has shifted to different kinds of PM. A historical summary of the PM NAAQS since 1971 is shown in Table 1. Originally the focus was on total suspended particulates (TSP), defined as anything that was collected on a fiber-glass filter in a High Volume Air Sampler. On average, these samplers would collect any particles in the air that were less than 30 - 40 µm in diameter, although this would vary with both wind speed and direction. However, from a health perspective, only particles that were small enough to be inhaled were of more concern. Consequently, in 1987 EPA promulgated the first inhalable PM NAAQS, which they called $PM_{10,}$ and is composed of particles that have a diameter of 10 µm or less. In the 1990s, additional attention became focused on fine particles which could be inhaled deeply into the lungs. Consequently, in 1997, EPA promulgated the first fine particle or $PM_{2.5}$ NAAQS. $PM_{2.5}$ are particles that have a diameter of 2.5 µm or less. The particles with a diameter larger than 2.5 µm but less than 10 µm are $PM_{10\text{-}2.5}$, which EPA has named the thoracic coarse fraction. In both the 2006 and 2012 PM reviews, EPA considered a separate NAAQS for these particles, but ended up concluding that the existing PM_{10} NAAQS covers them adequately. The June 2012 NPRM proposes to make the existing annual average $PM_{2.5}$ NAAQS more stringent and to retain the existing 24-hour NAAQS for $PM_{2.5}$ and PM_{10}. In addition, EPA has proposed a new secondary $PM_{2.5}$ NAAQS designed to protect visibility in urban areas.

It should be noted that another size classification that has recently received a great deal of attention is the ultrafine particles (UFP), which are a sub-class of $PM_{2.5}$ particles with diameters less than or equal to 0.1 µm. Since they are regulated as $PM_{2.5}$, EPA has not yet chosen to regulate these separately.

As a criteria pollutant, PM is unique in a very important way. The other criteria pollutants, carbon monoxide (CO), nitrogen dioxide (NO_2), sulfur dioxide (SO_2), ozone (O_3), and lead (Pb) are specific chemical compounds (or element in the case of lead) that produce relatively specific biological responses when inhaled. PM, on the other hand, consists of hundreds of chemical species with differing degrees of toxicity and health effects.

Final Rule	Primary/ Secondary	Indicator	Averaging Time	Level [1]	Form
1971 36 FR 8186 Apr 30, 1971	Primary	TSP [2]	24-hour	260 $\mu g/m^3$	Not to be exceeded more than once per year
			Annual	75 $\mu g/m^3$	Annual Average
	Secondary	TSP	24-hour	150 $\mu g/m^3$	Not to be exceeded more than once per year
1987 52 FR 24634 Jul 1, 1987	Primary and Secondary	PM_{10}	24-hour	150 $\mu g/m^3$	Not to be exceeded more than once per year on average over a 3-year period
			Annual	50 $\mu g/m^3$	Annual arithmetic mean, averaged over 3 years
1997 62 FR 38652 Jul 18, 1997	Primary and Secondary	$PM_{2.5}$	24-hour	65 $\mu g/m^3$	98th percentile, averaged over 3 years
			Annual	15.0 $\mu g/m^3$	Annual arithmetic mean, averaged over 3 years [3],[4]
		PM_{10}	24-hour	150 $\mu g/m^3$	Initially promulgated 99th percentile, averaged over 3 years; when 1997 standards for PM10 were vacated, the form of 1987 standards remained in place (not to be exceeded more than once per year on average over a 3-year period) [5]
			Annual	50 $\mu g/m^3$	Annual arithmetic mean, averaged over 3 years
2006 71 FR 61144 Oct 17, 2006	Primary and Secondary	$PM_{2.5}$	24-hour	35 $\mu g/m^3$	98th percentile, averaged over 3 years [6]
			Annual	15.0 $\mu g/m^3$	Annual arithmetic mean, averaged over 3 years [2], [7]
		PM_{10}	24-hour [8]	150 $\mu g/m^3$	Not to be exceeded more than once per year on average over a 3-year period

[1] Units of measure are micrograms per cubic meter of air ($\mu g/m^3$).
[2] TSP = total suspended particles.
[3]The level of the annual standard is defined to one decimal place (i.e., 15.0 $\mu g/m^3$) as determined by rounding. For example, a 3-year average annual mean of 15.04 $\mu g/m^3$ would round to 15.0 $\mu g/m^3$ and, thus, meet the annual standard and a 3-year average of 15.05 $\mu g/m^3$ would round to 15.1 $\mu g/m^3$ and, hence, violate the annual standard (40 CFR part 50 Appendix N).
[4]The level of the standard was to be compared to measurements made at sites that represent "community-wide air quality" recording the highest level, or, if specific requirements were satisfied, to average measurements from multiple community-wide air quality monitoring sites ("spatial averaging").
[5] See 69 FR 45592, July 30, 2004.
[6] The level of the 24-hour standard is defined as an integer (zero decimal places) as determined by rounding. For example, a 3-year average 98th percentile concentration of 35.49 $\mu g/m^3$ would round to 35 $\mu g/m^3$ and thus meet the 24-hour standard and a 3-year average of 35.50 $\mu g/m^3$ would round to 36 and, hence, violate the 24-hour standard (40 CFR part 50 Appendix N).
[7] The EPA tightened the constraints on the spatial averaging criteria by further limiting the conditions under which some areas may average measurements from multiple community-oriented monitors to determine compliance (see 71 FR 61165-61167).
[8] The EPA revoked the annual PM_{10} NAAQS in 2006.

Table 1: History of PM NAAQS from 1971 to present (U.S. EPA, 2012).

C. Health Assessments of Criteria Pollutants

In considering the establishment of NAAQSs, EPA has relied on three types of health effect studies: controlled human exposures, animal toxicology and epidemiology studies. In all NAAQS reviews prior to the 1996 PM review, EPA relied most heavily on controlled human exposures, which establish health effect endpoints as a function of exposure and demonstrate causality, and the toxicology studies which can provide insights as to the mode of the damage caused by an exposure. Epidemiology studies were used if they supported the findings in the other two types of studies because epidemiology studies can only identify statistical associations and cannot be used to demonstrate causality (cause-effect relationships). In the past, the discovery of a moderate association between a chemical substance (whether a pollutant or a drug) and a health endpoint simply meant that additional investigation was warranted, such as clinical and toxicological studies.

D. Health Assessments of Recent PM Reviews

1. 1996 PM Review

For the 1996 $PM_{2.5}$ review, EPA subordinated their reliance on human exposure and toxicological studies because they showed no evidence of effects at concentrations near the existing NAAQS. Instead, they relied primarily on epidemiology studies, which were finding very weak statistical associations between measures of PM, including PM_{10} and $PM_{2.5}$, and mortality at ambient concentrations well below the existing PM_{10} NAAQS. The fact that these studies conflicted with the results of human exposure and toxicology studies was ignored.

In the 1996 SP (U.S. EPA, 1996), EPA states on page VI-1: "the CD concludes that the overall consistency and coherence of the epidemiologic evidence suggests a likely causal role of ambient PM in contributing to adverse health effects." EPA did acknowledge that there were significant uncertainties in this conclusion. The following statements appeared in the SP in the section, "Summary of Key Uncertainties and Research Recommendations." They admitted that they have no mechanism to explain the epidemiology results.

> One of the most notable aspects of the available information on PM is the lack of demonstrated mechanisms that would explain the mortality and morbidity effects associated with PM at ambient levels reported in the epidemiological literature. The absence of such mechanistic information limits judgments about causality of effects and appropriate concentration-response models to apply in quantitatively estimating risks.

They admitted that there were likely serious measurement errors.

> Uncertainties and possible biases introduced by measurement error in the outdoor monitors, including both the error in the measurements themselves and the error introduced by using central monitors to estimate

> population exposure, contributes to difficulties in interpreting the epidemiological evidence.

EPA was also concerned that other pollutants present in the ambient air may be confounding the observed PM relationships.

> Inherent in epidemiological studies such as those cited in this review is the question as to whether or to what extent the observed effects attributed to PM exposures are confounded by other pollutants commonly occurring in community air, such as SO_2, ozone, NO_2, and CO. In particular, a number of authors conducting reanalyses of mortality studies within a given city, most notably for Philadelphia, have demonstrated that it may not be possible to separate individual effects of multiple pollutants when those pollutants are highly correlated within a given area.

Even with all the uncertainties, EPA felt that the weight of evidence leaned slightly more in favor of a real effect than not, but admitted they had no idea of what part of the fine PM was responsible for mortality.

> Although staff has concluded that it is more likely than not that fine fraction particles play a significant role in the reported health effects associations, identification of specific components and/or physical properties of fine particles which are associated with the reported effects is very important for both future reviews of the standards and in development of efficient and effective control strategies for reducing health risks.

Not knowing what constituent in the $PM_{2.5}$ is a causal agent is very problematic. EPA knew that PM is a mixture of many different substances with widely varying toxicities in controlled exposures. Regulating fine PM as if all the components are equally toxic is an assumption that cannot be supported based on the known toxic properties of the individual components. Many years of research on the toxicity of individual PM components demonstrate that the toxicity of PM components varies per unit of mass by a factor of at least 1,000.

EPA also admitted that they had no idea of what the dose-response function looks like or if there is a threshold below which there are no effects.

> Uncertainties in the shape of concentration-response relationships, most specifically whether linear or threshold models are more appropriate, significantly affects the confidence with which risks and risk reductions can be estimated.

So it is clear that in 1996, EPA recognized the uncertainties associated with the epidemiologically-based PM NAAQS and reducing them before the next PM review became a primary research focus of the Agency. Despite these realizations, EPA promulgated new annual and 24-hour $PM_{2.5}$ NAAQS based on the epidemiology findings.

2. 2006 PM Review

In spite of a massive effort by the Agency to reduce these uncertainties, they still existed in the 2006 review. However, in the final 2005 SP (U.S. EPA, 2005), EPA claimed they were making progress in some areas:

> For example, regarding the lack of demonstrated biological mechanisms, new evidence from toxicologic and controlled human exposure studies has provided information on an array of potential mechanisms for effects on the cardiac and respiratory systems, as discussed in Chapters 7 and 9 of the CD.

A key word in this quote is "potential." EPA had identified many possible mechanisms but no probable mechanism. However, buoyed by this apparent progress and the publication of numerous new epidemiological studies, many of which were funded by EPA, the 2004 CD (U.S. EPA, 2004) stated on page 9-48: "the epidemiological evidence continues to support likely causal associations between $PM_{2.5}$ and PM_{10} and both mortality and morbidity from cardiovascular and respiratory diseases, based on an assessment of strength, robustness, and consistency in results." They further added: "Epidemiologic studies suggest no evidence for clear thresholds in PM-mortality relationships within the range of ambient PM concentrations observed in these studies." However, EPA also was aware of other epidemiological studies that found health relationships with gaseous pollutants, so they tempered their conclusions somewhat later in the same chapter of the CD on page 7-79 when they stated: "A growing body of evidence both from epidemiologic and toxicologic studies also supports the general conclusion that $PM_{2.5}$ (or one or more $PM_{2.5}$ components), acting alone and/or in combination with gaseous co-pollutants, are likely causally related to cardiovascular and respiratory mortality and morbidity."

However, in the SP, EPA did highlight other remaining uncertainties and admitted that they were unusually large: "Staff believes it is important to continue to highlight the unusually large uncertainties associated with establishing standards for PM relative to other single component pollutants for which NAAQS have been set."

No progress had been made identifying a responsible agent (or agents) in the $PM_{2.5}$ and EPA admitted that their current strategy of treating all $PM_{2.5}$ equally may not be reducing health risks.

> Identification of specific components, properties, and sources of fine particles that are linked with health effects remains an important research need. Available evidence provides no basis for expecting that any one component would be solely responsible for $PM_{2.5}$-related effects, but it is likely that some components are more closely linked with specific effects than others. Continued source characterization, exposure, epidemiologic, and toxicologic research is needed to help identify components, characteristics, or sources of particles that may be more closely linked with various specific effects to aid in our understanding of causal agents

> and in the development of efficient and effective control strategies for reducing health risks.

They were still concerned about confounders in the pollution mix:

> The relationship between PM and other air pollutants in causing health effects remains an important question in reducing public health risk from air pollution. Numerous new analyses have indicated that associations found between PM and adverse health effects are not simply reflecting actual associations with some other pollutant. However, effects have been found with the gaseous co-pollutants, and it is possible that pollutants may interact or modify effects of one another. Further understanding of the sources, exposures, and effects of PM and other air pollutants can assist in the design of effective strategies for public health protection.

In addition, a new uncertainty caught the Agency's attention:

> Methodological issues in epidemiologic studies were discussed at length in the previous review, and it appeared at the time that the epidemiologic study results were not greatly affected by selection of differing statistical approaches or methods of controlling for other variables, such as weather. However, investigation of recently discovered questions on the use of generalized additive models in time-series epidemiologic studies has again raised model specification issues. While reanalyses of studies using different modeling approaches generally did not result in substantial differences in model results, some studies showed marked sensitivity of the PM effect estimate to different methods of adjusting for weather variables. There remains a need for further study on the selection of appropriate modeling strategies and appropriate methods to control for time-varying factors, such as temperature.

This last uncertainty has become known as the "model selection bias" issue and it is discussed in depth later in this report.

Despite these uncertainties, EPA used the epidemiology results as a basis for significantly lowering the 24-hour $PM_{2.5}$ NAAQS.

3. The 2008 Ozone, 2010 Nitrogen Dioxide and Sulfur Dioxide Reviews

Up to this point in time, the PM reviews were the only ones that were driven by epidemiology studies. In the 1997 review, EPA dismissed epidemiological studies that suggested that ozone, nitrogen dioxide and/or sulfur dioxide, not PM, were associated with premature mortality as being biologically implausible. However, the quote on the previous page from the 2006 PM SP, "However, effects have been found with the gaseous co-pollutants..." was foreshadowing a change in the Agency's tactics. Epidemiology studies that had been previously dismissed and new studies, some funded by EPA, were now being considered in the other NAAQS reviews.

Consequently, epidemiology played a central role in the 2008 ozone review and the 2010 reviews for both sulfur dioxide and nitrogen dioxide. As a result, EPA promulgated a new, very stringent eight-hour NAAQS for ozone, and new, very stringent one-hour NAAQS for nitrogen dioxide and sulfur dioxide. In one surprising outcome, EPA dismissed the epidemiology results that implicated CO to a variety of health effects as being inconclusive, and retained the existing CO NAAQS in 2011.

4. 2012 PM Review

Since the 1997 Review, fifteen years of research focused on finding the PM "smoking gun" have transpired and the uncertainties that existed in 1997 still exist. Unfortunately, EPA no longer acknowledges them and has dropped most of the cautionary language out of their 2009 ISA (U.S. EPA, 2009). These are the conclusions from that document (emphasis is EPA's):

> Together, the collective **evidence is sufficient to conclude that a causal relationship exists between short-term $PM_{2.5}$ exposures and cardiovascular effects**.
>
> Therefore, the evidence is sufficient to conclude that a **causal relationship is likely to exist between short-term $PM_{2.5}$ exposures and respiratory effects.**
>
> Collectively, the epidemiologic evidence is sufficient to conclude that **a causal relationship exists between short-term exposure to $PM_{2.5}$ and mortality.**
>
> Based on the above findings, the epidemiologic and toxicological evidence is **sufficient to infer a causal relationship between long-term $PM_{2.5}$ exposures and cardiovascular effects.**
>
> **Collectively, the evidence is sufficient to conclude that the relationship between long-term $PM_{2.5}$ exposure and respiratory effects is likely to be causal.**
>
> Collectively, the evidence is **sufficient to conclude that the relationship between long-term $PM_{2.5}$ exposures and mortality is causal.**

Some remaining uncertainties are acknowledged in the latest PA (U.S. EPA, 2011).

> The uncertainties and limitations that remain in the review of the primary fine particle standards are primarily related to understanding the range of ambient concentrations over which we continue to have confidence in the health effects observed in the epidemiological studies, as well as the extent to which the heterogeneity observed in the epidemiological evidence is

> related to differences in the ambient fine particle mixture and/or exposure-related factors. In addition, uncertainties remain in more fully understanding the role of $PM_{2.5}$ in relationship to the roles of gaseous co-pollutants within complex ambient mixtures.

The uncertainties that EPA acknowledges are, in their mind, minor inconveniences that hinder them from making conclusive statements with one hundred percent certainty. Therefore, they used the epidemiology studies as a basis for proposing a new, even more stringent annual $PM_{2.5}$ NAAQS in June 2012 (77 FR 38943) that, if implemented, will be the new controlling standard for $PM_{2.5}$. The range of the proposed NAAQS is 12 - 13 $\mu g/m^3$.

As we will show in subsequent sections of this report, the uncertainties still exist; they are not only significant, but they are game changers and they are more important than ever because EPA is now using the epidemiology results to drive all criteria pollutant NAAQS towards zero.

III. Epidemiology Study Results

A. What EPA Says They Show

In the 2009 ISA EPA makes numerous claims about cause-effect relationships between PM components ($PM_{2.5,}$ $PM_{10-2.5}$ and UFPs) and a variety of health effects that include mortality, respiratory and cardiovascular effects, reproductive and developmental effects and cancer, mutagenicity and genotoxicity. A listing of the effects that EPA evaluated, both short-term and long-term, along with EPA's causality determination for each effect is presented in Table 2-6 of the ISA. Rather than discussing all of EPA's interpretations on causality, the discussion below will focus only on the causal relationships that have driven EPA to propose the recent NAAQSs for not only $PM_{2.5}$, but also for ozone, sulfur dioxide and nitrogen dioxide. To further focus this discussion, we will limit it to short-term and long-term mortality studies because the same methodologies and same type of epidemiology studies are used to derive their causality conclusions on cardiovascular and respiratory morbidity. Therefore, the two conclusions of most concern are:

> Collectively, the epidemiologic evidence is sufficient to conclude that **a causal relationship exists between short-term exposure to $PM_{2.5}$ and mortality.**

and,

> Collectively, the evidence is **sufficient to conclude that the relationship between long-term $PM_{2.5}$ exposures and mortality is causal.**

The June 29, 2012 NPRM describes the process and the enormous amount of material EPA considered before they came to their conclusions (77 FR 38900):

> As discussed below, a substantial amount of new research has been conducted since the close of the science assessment in the last review of

> the $PM_{2.5}$ NAAQS (U.S. EPA, 2004), with important new information coming from epidemiological studies, in particular. This body of evidence includes hundreds of new epidemiological studies conducted in many countries around the world.

and,

> The newly available research studies as well as the earlier body of scientific evidence presented and assessed in the Integrated Science Assessment have undergone intensive scrutiny through multiple layers of peer review and opportunities for public review and comment. In developing this proposed rule, the EPA has drawn upon an integrative synthesis of the entire body of evidence between exposure to ambient fine particles and a broad range of health endpoints (U.S. EPA, 2009, Chapters 2, 4, 5, 6, 7, and 8) focusing on those health endpoints for which the Integrated Science Assessment concludes that there is a *causal* or *likely causal relationship* with long- or short-term $PM_{2.5}$ exposures.

The studies they relied on for short-term mortality are displayed in Figure 6-27 of the ISA and are summarized in section 6.5.3.1:

> $PM_{2.5}$ risk estimates were found to be consistently positive, and slightly larger than those reported for PM_{10} for all-cause, and respiratory- and cardiovascular-related mortality. The risk estimates for all-cause (nonaccidental) mortality ranged from 0.29% (Dominici et al., 2007) to 1.21% (Franklin et al., 2007) per 10 $\mu g/m^3$ increase in $PM_{2.5}$. These associations were consistently observed at lag 1[2] and lag 0-1, which have been confirmed through extensive analyses in PM_{10}-mortality studies. Cardiovascular-related mortality risk estimates were found to be similar to those for all-cause mortality; whereas, the risk estimates for respiratory-related mortality were consistently larger: 1.01% (Franklin et al., 2007) to 2.2% (Ostro et al., 2006) using the same lag (i.e., lag 1 and lag 0-1) and averaging indices.

Similarly, the studies they relied on for long-term mortality are summarized in Figures 7-6 and 7-7 of the ISA and in section 7.6.5.1:

> The recent evidence is largely consistent with past studies, further supporting the evidence of associations between long-term $PM_{2.5}$ exposure and increased risk of human mortality (Section 7.6) in areas with mean concentrations from 13.2 to 29 $\mu g/m^3$ (Figure 7-7). New evidence from the Six Cities cohort study shows a relatively large risk estimate for reduced mortality risk with decreases in $PM_{2.5}$ (Laden et al., 2006). The results of new analyses from the Six Cities cohort and the ACS study in Los

[2] Lag refers to the number of days between the exposure and the occurrence of the health effect. A lag of 1 means the effect occurred one day after the exposure.

> Angeles suggest that previous and current studies may have underestimated the magnitude of the association (Jerrett et al., 2005). With regard to mortality by cause-of-death, recent ACS analyses indicate that cardiovascular mortality primarily accounts for the total mortality association with $PM_{2.5}$ among adults, and not respiratory mortality. The recent WHI cohort study shows even higher cardiovascular risks per $\mu g/m^3$ than found in the ACS study, but this is likely due to the fact that the study included only post-menopausal women without pre-existing cardiovascular disease (Miller et al., 2007). There is additional evidence for an association between $PM_{2.5}$ exposure and lung cancer mortality (Section 7.5.1.1).

Thus, EPA presents a coherent and impressive argument to support their conclusions of causal associations between ambient $PM_{2.5}$ concentrations and both short- and long-term mortality. Unfortunately, as shown in the next sections of this report, the results have been cherry-picked and conflicting evidence has either been ignored or misrepresented by the Agency.

B. Another Perspective on the Epidemiology Studies

Coherence of results between epidemiology and toxicology and dosimetry studies provides a reality check for the epidemiology results. Although the CD (US EPA, 2004) and ISA (US EPA, 2009) documents include extensive discussion of the deposition of particles in the respiratory tract (dosimetry) and the results of controlled human and animal exposures (toxicology), the June 2012 NPRM (77 FR 38890-39055) includes little or no mention of these findings. Rather the NPRM relies on the epidemiological discussions in the ISA and EPA's risk assessment (US EPA, 2010), that is based on a series of questionable assumptions. Since the findings from dosimetry and toxicology are highly relevant to the judgment as to whether particles are actually causing health effects at current ambient levels, those findings are summarized and discussed in the following section.

1. Dosimetry

Information on the dosimetry of particles, that is the deposition, clearance, and retention of particles within the respiratory tract, is critical to understanding the health effects of inhaled particles. The information is critical because the cause of a biological response to PM is due to the dose deposited at the target site rather than the external exposure. The respiratory tract can be divided into three regions, the extra-thoracic (mouth and nose), the tracheobronchial (the trachea and a series of conducting airways that branch for a number of generations), and the alveolar (the smallest branching airways and the small sacs where air exchange with the blood takes place). The body has a number of mechanisms to deal with inhaled particles. These include the flow of mucus towards the pharynx where material is swallowed and enters the GI tract, dissolution into the blood stream, and ingestion by cells called macrophages that are part of the body's immune system and detoxify and remove unwanted material. Thus, the fate of a particle in the body and its potential effects will vary based upon where it is deposited, its degree of solubility, and other factors. Dosimetric information is critical for comparing effects noted in animal toxicology experiments or controlled human studies with the doses implicated by epidemiological studies.

Information from dosimetry can answer two key questions. Both these questions arise because of EPA's concerns that health effect associations are most likely related to the $PM_{2.5}$ fraction and to combustion-related chemical constituents. The first question is whether the focus on fine particles is justified by dosimetry? The second question is whether the doses of fine particles to target tissue in a 24-hour period or over a lifetime are high enough to cause the effects implied by the statistical associations?

The 2004 CD included a plot of fractional deposition versus particle size (Figure 6-19) indicating that there is substantial overlap between the deposition pattern of fine and coarse particles in the respiratory tract. This is in agreement with Snipes et al., 1997, who showed that particle deposition per unit surface area decreases by orders of magnitude from the extrathoracic to the tracheobronchial and to the alveolar regions. In addition, Snipes et al. showed that coarse and fine particles are deposited in both the tracheobronchial and alveolar regions. Snipes et al. concluded that their modeling "demonstrated significant thoracic deposition of environmental aerosol particles larger than those collected in a $PM_{2.5}$ sampler" and recommended that PM_{10} rather than $PM_{2.5}$ would be a "good indicator of potential health effects." Thus, the answer to the first question is that, from dosimetry, there is no reason to focus concern on fine particles to the exclusion of coarse particles.

There are several studies that address the second question. Snipes et al., 1997 estimated that the mass of particles deposited per unit of alveolar–interstitial tissue in humans inhaling particle concentrations as high as 50 $\mu g/m^3$ for 24-hours was only in fractions of nanograms (10^{-9} gram) of particles per square centimeter. Vostal (2000) extended the Snipes et al. calculations to include individual $PM_{2.5}$ components using speciated data from U.S. metropolitan areas in Texas. For individual particle components that are of interest as potential causal agents, much lower deposition levels were found. For example, sulfate deposits were only in the range of picograms (10^{-12} gram) per square cm of alveolar surface, and levels of elemental carbon, iron or trace elements were not higher than a fraction of a picogram per square cm of surface. For toxic metals, suggested as a probable cause of fine particle toxicity, the estimated 24-hour deposition levels were extremely low, not exceeding tens of femtograms (10^{-15} gram) per square cm of alveolar-interstitial surface.

Similar results to those of Snipes et al. and Vostal are reported by Winter-Sorkina and Cassee (2002). Thus, there is consistent information on the very low magnitude of deposited doses implicated by acute time-series epidemiology.

Other factors can influence the distribution of deposits such as gender, age, or non-homogenous distributions of particle loads in diseased lungs. However, these factors can change the deposition levels only by small multiples and will not change the low order of magnitude of the daily doses. In addition, when one considers the solubility and systemic distribution of PM in the organism, it is clear that even if the total deposited amounts are dissolved (which is not very probable) and distributed in the organism, the total systemic dose will be greatly diluted.

The challenge for toxicology is to explain how such low doses of particles can be causing the health effects implied by the epidemiological associations. The toxicological evidence is discussed in the next section.

2. Toxicology of Particulate Matter in Humans and Experimental Animals

One of the major uncertainties EPA acknowledged during the 1996 review of PM NAAQS was the lack of demonstrated mechanisms that would explain the mortality and morbidity effects implied by the epidemiological associations. A review of the toxicology material in the 2009 ISA reveals that, despite over a decade of expanded and focused research, there are still no data from controlled studies that indicate how anthropogenic PM at current ambient levels is causing the mortality and morbidity effects implied by the epidemiological associations EPA relies on.

The experimental techniques typically used include inhalation, intratracheal instillation, and in vitro exposures of cells in solution or suspension. However, inhalation is the only realistic exposure regime for identifying a toxic dose. Direct instillation of material to the lung tissue leads to a different distribution of the material in the respiratory tract compared to inhalation and can overwhelm the body's natural defense mechanisms. In vitro experiments use unrealistically high exposures of live cells to gain information concerning mechanisms of action. Thus, only inhalation experiments are appropriate for estimating human risk.

The use of high doses and non-physiological means of exposure complicate the interpretation of the many studies reviewed in the ISA. In fact, the 2004 CD acknowledged that "one overarching issue in the interpretation of toxicology study results is the relevance of findings from experimental human or animal studies using controlled exposure/dose concentrations that are high relative to the much lower ambient pollutant exposure levels that apply within the context of pertinent epidemiology studies" (page 7- 205).

The ISA reviews literally hundreds of references. However, many of the new studies were designed to address the question of the kinds of effects that PM mixtures or PM constituents can cause and few address the more important issue of what the dose-response is for controlled PM exposures. The 2004 CD noted "An important caveat in interpretation of the toxicological data is that the high doses used in many of the studies may produce different effects on the lung than inhalation exposures at lower ambient concentrations" (Page 7-206). It went on to note that high experimental doses may activate cells and pathways entirely different from those activated by more realistic doses. This is a well-known phenomenon in toxicology.

Toxicology is known as the science of poisons, where a poison may be any substance which when acting directly through its inherent chemical properties is capable of destroying or seriously endangering life. Any substance, even food and water, may be harmful if absorbed in excessive amounts. The dose determines whether or not injury will occur, requiring the toxicologist to pay careful attention to the quantitative measurement of both dosage and effect, before the delivered dose is declared as "harmful." In fact, in the testing of single chemical compounds, large spectrums of biological manifestations as well as the magnitude of the effects are examined. These manifestations range from completely negative responses to harmful organ damage or to generalized systemic responses, which are all predetermined by the size of the dose. The dose makes the poison.

In discussing respiratory effects, the 2004 CD indicated that PM effects that vary with chemical and physical characteristics have been extensively studied for over 30 years. (Page 7-85) It also stated that the data provide little basis for concluding that specific PM constituents have substantial respiratory effects at current ambient levels. This substantial body of information is

routinely used to establish chemical-specific standards that are used in risk assessment in occupational and other environmental settings. In fact, EPA uses chemical-specific standards in the risk assessment of emissions from Superfund sites and for assessing the risk from hazardous air pollutants. The standards for various PM materials clearly show that the relative toxicity of different PM species as measured per unit mass varies by over a factor of 1,000.

Valberg (2004) used the chemical-specific, dose-response data typically used in U.S. EPA human health risk-assessments to evaluate the risk associated with a mixture of 27 separate chemical constituents typical of ambient PM with a total $PM_{2.5}$ concentration of 15 $\mu g/m^3$. The assessments rely on established, no-effect thresholds for noncancer health endpoints. Valberg found that the chemicals identified as constituents of ambient PM are present at concentrations considerably below the regulatory thresholds (for which no adverse health effects are anticipated for a lifetime exposure) used in risk assessment. From the perspective of risk assessment, Valberg concluded that exposure to the concentrations of chemicals in ambient PM (e.g., sulfate, nitrate, and 25 other constituents) cannot be expected to cause death. Hence, he noted that the health effects attributed to ambient PM in the NAAQS review appear to be at odds with what would be predicted from a standard U.S. EPA health-risk assessment for PM chemicals. Valberg discusses four possible explanations for this paradox: 1) the toxicity of ambient PM is unrelated to its chemical constituents, 2) PM mass concentration is not the causal factor in the reported associations, 3) the mixtures of chemicals in ambient PM are vastly more toxic than the sum of individual components, or 4) a small portion of the general population are vastly more sensitive to certain PM chemicals than reflected in the EPA toxicity factors. As shown below, however, a more likely explanation of this paradox is model selection bias, confounding and exposure uncertainty. The EPA rulemaking materials (ISA, PA, NPRM) are silent on the existence of this paradox, much less on the possible explanations.

In the previous review, CASAC specifically commented that the CD "…must make it clear that there is a large data base that indicates that PM is markedly variable in its toxic potency." In addition, a blue-ribbon Committee set up under the National Research Council to advise the Agency on PM research concluded that the current mass-based NAAQS that implicitly assumes that all $PM_{2.5}$ particles have the same toxicity per unit mass, irrespective of chemical composition "greatly oversimplifies complex biological phenomena" (NRC, 2004). Thus, the assumption that all PM is equally toxic cannot be supported.

Historically, our knowledge of the health risk from different kinds of particles came from occupational studies and animal and human studies of controlled exposures that identified specific health problems that were caused by specific kinds of particles. For example, high concentrations of quartz and silica cause respiratory disease; asbestos can cause lung cancer and mesothelioma; coal dust can cause black lung, but not cancer; specific metals like nickel, cadmium, and beryllium can cause cancer or respiratory disease; and certain biological particles like molds and pollens can cause allergic responses. Many other particles were historically regulated under the general category of nuisance dust.

For some of the major components of atmospheric PM, such as sulfates and nitrates, extensive studies with animals and humans did not show effects until the concentrations were many times current ambient levels. For example, Schlesinger and Cassee (2003) concluded, in a review of

the toxicology of sulfates and nitrates, that:

> …the currently available toxicological database does not support a role for secondary inorganic aerosols in adverse health outcomes noted in epidemiological studies, in that levels of these particles, and specifically the most toxicologically potent acid species, needed to produce any effect in controlled studies are well above those found in ambient air in the United States.

Another review by Reiss et al. (2007) concluded that:

> In total, the epidemiologic and toxicologic evidence provide little or no support for a causal association of PM sulfate and health risk at ambient concentrations. For nitrate-containing PM, virtually no epidemiological data exist. Limited toxicological evidence does not support a causal association between particulate nitrate compounds and excess health risks.

There is another recent major toxicological study of power plant emissions that confirms the limited toxic potential of current exposures to sulfate and nitrate. In a series of papers in *Inhalation Toxicology*, Godleski et al. (2011) report on a toxicological evaluation of coal-fired power plant emissions. Godleski et al. indicate:

> The toxicological evaluation of realistic emissions of source aerosols (TERESA) study involved withdrawal of emissions directly from the stacks of three coal-fired power plants. The emissions were aged and photochemically transformed to simulate downwind power plant plume processing. In order to carry out these studies in the field at the power plants, mobile laboratories were constructed which included reaction chambers, instrumentation for characterization, and animal exposure chambers.

Kang et al. (2011) indicate:

> Test atmospheres developed for toxicological experiments included scenarios to simulate a sequence of atmospheric reactions that can occur in a plume: (1) primary emissions only; (2) H_2SO_4 aerosol from oxidation of SO_2; (3) H_2SO_4 aerosol neutralized by gas-phase NH_3; (4) neutralized H_2SO_4 with secondary organic aerosol (SOA) formed by the reaction of α-pinene with O_3; and (5) three control scenarios excluding primary particles. The aged particle mass concentrations varied significantly from 43.8 to 257.1 $\mu g/m^3$ with respect to scenario and power plant.

Godleski et al. (2011) note:

> Toxicological outcomes were evaluated in Sprague-Dawley rats exposed to different emission scenarios. Breathing pattern, pulmonary

> inflammatory responses, in vivo pulmonary and cardiac chemiluminescence and cardiac response in a model of acute myocardial infarction were assessed.

Godleski et al. (2011) in summarizing the study note:

> The results showed no response or relatively mild responses to the inhaled aerosols studied; complex scenarios which included oxidized emissions and α-pinene to simulate biogenic secondary organic aerosol tended to induce more statistically significant responses than scenarios of oxidized and non-oxidized emissions alone. Relating adverse effects to specific components did not consistently identify a toxic constituent. These findings are consistent with most of the previously published studies using pure compounds to model secondary power plant emissions…

Thus, there was no evidence from the TERESA study to indicate that sulfate from power plants is any more toxic than previously understood from published studies using pure compounds.

One of the methods that has been used to study ambient PM in a semi-controlled way is to expose humans or animals to ambient air particles that have been concentrated by a factor of from 6 to 20. In studies with concentrated ambient air particles (CAPs), no consistent pattern of inflammatory changes has emerged. The few statistically significant changes that have been reported are small, transient, and within the normal physiologic range. It is not clear if these small changes are real changes that are not consistent because of the varying composition of the PM or if they are changes within the normal range solely due to chance.

A Health Effects Institute report (HEI, 2003a) on CAPs, diesel exhaust exposures and inflammation notes "a consistent pattern of inflammation after exposure to a variety of PM mixtures in many studies has not emerged to date." The synopsis notes for example that "many markers of inflammation were studied but few changed; of those that changed, the magnitude of the change was modest." It was also noted "because so few markers of inflammation changed in the current studies, it is possible that these changes occurred by chance." Thus, with exposures to elevated concentrations of concentrated ambient particles there are, at the most, small transient changes that are within the normal physiologic range and not of any clinical significance. Such changes cannot explain the epidemiologic associations.

Godleski et al. (2011) compared their results to those from exposures to concentrated ambient air particles (CAPs) concluding "In general, only the most complex scenarios approached, but did not equal or exceed the reported toxicity of inhaled CAPs." Thus, studies of the toxicity of concentrated ambient air particles and concentrated coal power plant emissions cannot explain the epidemiologic associations.

Although the 2009 ISA summarizes many relevant toxicological studies, evaluating potential mechanisms by which PM might affect respiratory or cardiovascular health, there are still no demonstrated mechanisms by which anthropogenic PM could cause the serious health effects inferred from the epidemiologic associations in the literature.

While there have been biological effects from many PM components shown at unrealistically high doses, or with un-physiologic methods of exposure, neither the presence nor the plausibility of effects at relevant human doses has been demonstrated. At high doses, in comparative toxicology studies, metals and bioaerosols[3] seem to have the most toxicity under what are admittedly unrealistic test conditions. Thus, among the various constituents of ambient PM, they are components that deserve high priority for study at relevant ambient doses. In fact, certain bioaerosols are the only PM constituents that have been shown to have health effects at or near current exposure levels. The 2004 CD acknowledged that pollens, fungal spores, bacteria, and viruses have all been shown to cause or contribute to adverse health effects. (Page7-220)

In evaluating the toxicological literature, it is important to note that almost every published study reports some positive effect. This arises because investigators do not like to submit and editors do not like to publish negative studies. However, when the full pattern of results is evaluated, along with the fact that studies evaluate many possible endpoints including a wide variety of sub-clinical biomarkers, the toxicological evidence for PM health effects due to anthropogenic constituents at ambient concentrations is very weak.

For example, the ISA acknowledges that respiratory symptoms have not been reported following controlled exposures to $PM_{2.5}$ among healthy or health-compromised adults. Similarly, pulmonary function in humans is not affected. Whereas some changes in biomarkers of pulmonary inflammation have been reported from CAPs or other controlled studies in humans or animals, the ISA indicates that the response appears to vary significantly depending on the source and composition of the particles. It is important to note that the animal exposures where mild responses are reported are typically greater than 1 mg/m^3 (1,000 $\mu g/m^3$) which is hardly relevant to whether the annual NAAQS should be 12 or 13 or 15 $\mu g/m^3$.

To explain potential cardiovascular PM effects, systemic inflammation has been studied in toxicological studies. However, the ISA concludes with regard to human exposures that "New studies involving controlled exposures to various particle types have provided limited and inconsistent evidence of a PM-induced increase in markers of systemic inflammation." Increased blood coagulability secondary to lung inflammation is another potential mechanism that has been postulated and studied. However, the 2004 CD indicated that the results are mixed and inconsistent. The 2009 ISA still refers to the data as inconsistent.

Effects on autonomic control of the heart and cardiovascular system are another potential mechanism under active study. However, the ISA also acknowledges inconsistent results for this potential mechanism. The findings of these new studies do not provide convincing evidence of an association between PM exposure and an increase in blood pressure or hypertension or of a consistent association with heart rate variability.

Despite the outpouring of toxicological studies evaluating many possible mechanisms by which $PM_{2.5}$ may cause harm, there is a lack of consistent findings of clinically-relevant PM effects at high concentrations. Thus, toxicology cannot explain how low concentrations can be causing death.

[3] Bioaerosols are PM that contain living organisms or were released from living organisms. They include: bacteria, viruses, fungal spores, plant pollen, dander, insect and plant fragments, etc.

3. Epidemiology Results Are Inconsistent with PM Being Causal

Clearly there have been historical air pollution episodes like the London and Donora, PA episodes where relationships between air pollution and acute health effects have been documented. In the U.S. today, levels of PM, even in cities with the highest concentrations are a fraction of what they were during these episodes. As a result, there is doubt in the minds of many air pollution professionals that today's levels of PM (PM_{10} or $PM_{2.5}$) in the U.S. are causing adverse health effects.

A close examination of the results of the epidemiology studies cited by EPA reveals that they do not produce consistent results. EPA relies on two types of epidemiology studies: "long-term," or "chronic," studies in which health effects in cities with different levels of air pollution are compared over long periods of times; and "time-series," "short-term" or "acute" studies in which health effects (e.g., deaths, emergency room visits) are compared within a city as air pollution levels fluctuate. Many of the recent short-term studies have also been "multi-city" studies as they compare the single-city time-series results with other cities and then pool these results to estimate regional or national averages.

a. NMMAPS

One reason for scientists' skepticism about whether PM is causing adverse health effects is that studies have consistently shown that a criteria pollutant that is associated with health effects in one area may not be associated with health effects in another area. The most striking example of this is from the multi-city National Morbidity Mortality Air Pollution Study (NMMAPS) (Dominici et al., 2003) which examined statistical associations between most criteria pollutants (PM_{10}, ozone, carbon monoxide, sulfur dioxide and nitrogen dioxide) in single-pollutant models, but made the results available for each of the 90 individual cities before pooling the time-series derived effect estimates to obtain a single national estimate. Only the lag 1 results will be discussed because they gave the maximum effects for all pollutants.

The results from NMMAPS are presented in Figures A1 - A10 and Table A1 in the attached Appendix A. Figures A1 - A5 show the risk estimates for each pollutant in increasing order by city. Figures A6 -A10 show the same results cluster by geographical region. The results show two important features found in the results of all the multi-city studies. First, the results exhibit significant heterogeneity (i.e., they were inconsistent) both across the country and within each geographical region. For PM_{10}, the ranges of risks are implausible and inconsistent with a causal PM_{10}/mortality relationship. The risks range from -3.4 to +3.0 with 63% of the cities displaying a positive statistical relationship between mortality and PM_{10}. Only 2 cities (New York and Oakland) have a statistically significant positive effect. On the other hand, 37% of the cities have a zero or negative relationship. Taken at face value, a negative effect would imply a biologically implausible protective effect (i.e. PM_{10} provides protection from mortality).

Second, the distribution of risks across all the cities is nearly the same for all the pollutants - more than half exhibit a positive risk and a quarter to nearly half show a zero or negative risk. It

seems irrational to single out PM_{10} as the causal agent when the results for the other criteria pollutants are nearly identical.

In NMMAPS, the investigators then take the individual city risk estimates and combine them using a Bayesian procedure to arrive at a single mean estimate of risk for the entire U.S. They also divide the U.S. into seven geographical sectors and compute a mean for each sector. For the entire U.S., they computed a statistically significant overall risk estimate of a 0.27% increase in mortality risk per 10 $\mu g/m^3$ increase in PM_{10}. In an excellent review and critique of air pollution epidemiology, Suresh Moolgavkar explains why such single estimates of risk based on Bayes analyses "has little meaning" (Moolgavkar, 2005).

In the geographic sector analyses, only the Northeast had a statistically positive risk estimate. While the risk estimates for the other geographic areas were positive, none were statistically significant.

b. Other Multi-City Studies

EPA's 2009 PM ISA correctly notes new multi-city studies that report major differences in PM associations as a function of geography and season. For example, all of the studies identified in the current PM ISA that have examined the PM-mortality relationship, in regards to geographic location within the U.S., have concluded that the effects are greater in the East compared to the West.

The NMMAPS analysis by season and region by Peng et al. (2005) which used updated mortality data from 1987-2000 in 100 cities, reported that summer was the only season for which the combined effect was statistically significant. An analysis by geographical regions showed a strong seasonal pattern in the Northeast with a peak in the summer and little seasonal variation in the southern regions of the country. The authors acknowledge that there are several possible explanations for their results. One obvious hypothesis is that the most toxic particles have a spring/summer maximum and are more prevalent in the Northeast. Another hypothesis mentioned by the authors is that there could be a seasonally varying bias from an, as yet, unidentified source.

The largest hospital admissions study also clearly shows differences in cardiovascular hospital admissions between East and West. The Dominici et al. (2006) study evaluated fine PM-hospital admissions associations for 204 U. S. urban counties with a population greater than 200,000 using 1999-2002 Medicare hospital admission data. The results are presented for a two-stage Bayesian analysis for various types of admissions and by region. Combined associations on the order of a 1 % increase in various cardiovascular or respiratory outcomes per 10 $\mu g/m^3$ increase in $PM_{2.5}$ are reported. However, there are issues that call into question the interpretation of this as an effect from generic fine PM.

The authors present results from seven separate regions as well as a comparison of the three western regions with the four eastern regions. There is a clear difference in the combined associations among the regions and particularly between the eastern and western regions. The combined association is positive for cardiovascular outcomes in the east, but negative in the

west, except for heart failure, which is positive in both areas. This is not consistent with an effect of generic $PM_{2.5}$ on cardiovascular hospital admissions and, indeed, the authors point out the need to shift the focus of research to identifying those characteristics of particles that determine their toxicity.

A similar spatial pattern exists in the chronic studies. The HEI-sponsored re-analysis of the Six-City and ACS studies (Krewski et al., 2000) showed that the increased risk was cardiovascular not respiratory, and there was significant spatial heterogeneity in the association, with no effect seen in western U. S. cities. In fact, a negative estimate of excess $PM_{2.5}$ mortality risk was found in the West. Krewski et al. also identified other patterns in the data including: SO_2 had a strong association with mortality, the PM all-cause mortality association was significantly reduced and became non-significant when SO_2 was added in a two pollutant model, and the increased mortality only occurred in the participants that had a high school education or less.

A recent analysis by Zeger et al. (2008) confirms the large spatial difference in chronic mortality association in a cohort of 13 million Medicare enrollees. Zeger et al. reported statistically significant results for the eastern and central United States that are in general agreement with previous publications, but Zeger et al. found no significant effect of $PM_{2.5}$ on mortality in the western United States. A caution in interpreting the Zeger study is that effect estimates for the Medicare cohort may be biased upward due to lack of adjustment for individual level risk factors.

c. APHENA

In October, 2009, the Health Effects Institute (HEI) published the results of the *Air Pollution and Health: A European and North American Approach (APHENA)* study (Katsouyanni and Samet, 2009). The APHENA project was designed to take advantage of the largest databases available. 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 European 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 was 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 method was best, 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 PM_{10}. In addition, subsets of these choices were also used to examine the effects of controlling for ozone.

The overall PM modeling results for the mortality models and the morbidity models are summarized in Tables 2 and 3, 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 PM_{10} 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 ($\approx 1/2$) 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 significant and positive relationship represents real causal relationship or if they are false positives that can occur by chance or by confounding.

For mortality, the strongest and most consistent significant relationships are observed for all cause and cardiovascular mortality, but only for the ≥ 75 years age group in Canada and Europe. Importantly, the signal is inconsistent in the U. S. as it is model dependent. For the younger age group, few models are significant except in Europe for all cause but not cardiovascular or respiratory. None of the three geographic areas show consistent significant positive model results for respiratory mortality. Further, none of the models in Canada produce significant results for respiratory mortality.

The models also show mixed results for the hospital admissions. The most consistent significant positive signal is seen for cardiovascular admissions in the U. S. and to a slightly less degree in Europe. However, none of the model formulations produce significant results in Canada. No consistent results are seen for respiratory admissions anywhere. They are strongly model dependent.

The above results from the APHENA study demonstrate the importance of model selection. However, APHENA did not undertake an exhaustive, comprehensive analysis of model selection as they include a limited number of model choices and only considered two pollutants, PM_{10} and ozone. The importance of model selection will be discussed in section IV.

While there are positive and significant combined associations for some models and for some endpoints and for some geographic areas, the overall pattern of associations in the large APHENA study is mixed and inconsistent. The overall pattern is not what one would expect if PM health effect associations have a real physiological basis. For example, it is not logical that PM would be causing cardiovascular hospital admissions in the U. S. but not in Canada. It is not logical that PM would have a strong cardiovascular mortality signal in Canada but not in the U.S.

It should be noted that APHENA conducted the identical analyses with ozone data and the results showed a similar pattern of mixed and inconsistent results.

Cause of Death	Canada	Europe	United States
All Cause – all ages	8/8	18/24	15/24
≥ 75 yrs	8/8	21/24	15/24
< 75 yrs	4/8	16/24	8/24
All Cause ozone controlled – all ages	8/8	16/16	9/16
≥ 75 yrs	8/8	13/16	10/16
< 75 yrs	0/8	13/16	4/16
Cardiovascular – ≥ 75 yrs	8/8	19/24	16/24
< 75 yrs	0/8	8/24	2/24
Cardiovascular –ozone controlled ≥ 75yrs	7/8	16/16	10/16
< 75 yrs	0/8	6/16	2/16
Respiratory – all ages	0/8	11/24	7/24
≥ 75 yrs	0/8	11/24	4/24
Respiratory – ozone controlled – all ages	0/8	7/16	3/16
≥ 75 yrs	0/8	7/16	3/16

Table 2: APHENA modeling results for mortality. The numerators represent the number of models that showed a positive and statistically significant relationship between PM_{10} and mortality while the denominator is the total number of models run.

Type of Admission	Canada	Europe	United States
Respiratory	2/8	16/24	9/24
Respiratory – ozone controlled	0/8	10/16	10/16
Cardiovascular	0/8	20/24	24/24
Cardiovascular – ozone controlled	0/8	12/16	16/16

Table 3: 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 PM_{10} and admissions while the denominator is the total number of models run.

IV. Why the Epidemiological Studies Find Inconsistent Results

A. Publication Bias

Publication bias is another major issue in assessing the epidemiological literature. Publication bias occurs because authors are inclined to mine the data for positive results and editors more likely to publish a paper with positive findings. Consequently, there will be more papers in the literature that show positive epidemiology results than those that show negative results. Thus any meta-analysis performed on the air pollution epidemiology literature uses biased inputs and the results are thus biased. The commentary by Goodman (2005) concerning meta-analyses is particularly insightful. He noted a factor of at least three difference between the results of ozone meta-analyses and the NMMAPS individual city results which are not affected by publication bias. Goodman concludes that the implications of an EPA-sponsored exercise of funding three separate meta-analyses "go far beyond the question of the ozone mortality effect." He cautions that "depending on published single-estimate, single-site analyses are an invitation to bias." He notes that "the most plausible explanation is the one suggested by the authors, that investigators tend to report, if not believe, the analysis that produces the strongest signal; and in each single-site analysis, there are innumerable model choices that affect the estimated strength of that signal." A separate review by a panel of ten knowledgeable scientists concluded that "taken together, the meta-analyses provide evidence of a disturbingly large publication bias and model selection bias" (Rochester Conference Report, 2007).

Similarly, Anderson et al., 2005 concluded that publication bias is present in single-city time series studies of ambient particles. After correcting for publication bias, they still report a positive association. However, they also note that the regression estimates from the multi-city studies (which are not prone to publication bias) and the corrected single-city studies are approximately half of the mortality estimates of the mid-1990's, that the correction for publication bias may not be complete, and that differential selection of positive lags may also inflate estimates.

Thus, publication bias is a major concern inflating the size of any potential effect. As EPA has reviewed other criteria pollutants, the Agency has acknowledged (US EPA, 2008a, b) that the summary of health effects evidence is vulnerable to the errors of publication bias and multiple

testing. The only reference in the PM ISA to publication bias is found on page 6-4 in a discussion of the heart rate variability findings. This ignores the fact that there is now substantial evidence that publication bias inflates the apparent magnitude and consistency of air pollution health effects in single-city studies.

B. Confounding

1. Short-Term Studies

All epidemiology studies must deal with the issue of confounding. Although the definition is somewhat technical,[4] it can be illustrated by means of well-known example. Suppose a statistical association between alcohol consumption and oral cancer is observed. It cannot be concluded, on the basis of this statistical association alone, that drinking alcohol causes oral cancer because this association may actually be reflecting the fact that smokers can also be consumers of alcohol. Epidemiology studies indicate that smoking is a strong risk factor for oral cancer and a study of alcohol and oral cancer that did not adequately control "confounding" by smoking would lead to biased conclusions regarding the association of alcohol consumption with oral cancer.

The ambient air in urban areas contains trace amounts of hundreds of chemical species both in the gas and particulate phase. Most epidemiology studies only focus on PM mass, but PM contains measureable amounts of nearly every element that exists in the earth's crust. Individual elements can exist as different chemical compounds. Therefore, there are hundreds of potential confounders in the air and only a tiny fraction of them are measured. Consequently, in a study of any one component of air pollution, other components that may be associated with health impacts must be controlled. Very few studies do this for even the ones that are measured. This means that the potential for confounding by other substances in the atmosphere can never be eliminated.

In all air pollution epidemiology, weather is also an obvious confounder. In addition, other temporal effects such as season, cyclic diseases, and day-of-the-week patterns must be controlled for. In a discussion of this subject, a Special Panel of HEI's Health Review Committee (Special Panel of the Health Effects Review Committee, 2003) noted:

> Neither the appropriate degree of control for time in these time-series analyses, nor the appropriate specification of the effects of weather, has been determined. This awareness introduces an element of uncertainty into the time-series studies that has not been widely appreciated previously. At this time, in the absence of adequate biological understanding of the time course of PM and weather effects and their interactions, the Panel recommends exploration of the sensitivity of these studies to a wider range of alternative degrees of smoothing and to alternative specifications of weather variables in time-series models.

[4] A confounder is an extraneous variable that correlates with both the dependent and independent variable. Such a relationship is termed a spurious relationship. In the case of a risk assessment, it is important to control for confounding to isolate the risk of a particular hazard.

In other words, it is widely known that weather and temporal confounders must be controlled, but the correct method to do so is not known.

Many short-term, time-series studies of air pollution report associations between fine PM and various measures of human health, such as the numbers of daily deaths and hospital admissions, in single pollutant analyses. However, when possible confounding by other pollutants is explicitly addressed, many of the studies find no association between fine PM and measures of human health. A good example of this is NMMAPS (Dominici et al., 2003) which showed that in single pollutant models each criteria pollutant had a statistically positive association with mortality. When two or more pollutant models were used, the coefficients were attenuated and, in most cases, lost statistical significance.

A study by Vedal et al. (2004) conducted in Vancouver reported a statistically significant association between sulfur dioxide and cardiac arrhythmias in a small subset of the original study population, but no statistically significant associations with either PM or ozone. This indicates that even in areas with low concentrations of SO_2, the SO_2 must be controlled as a possible confounder. This finding leads to a dilemma: if PM is responsible for adverse health outcomes, why are associations seen in these studies with SO_2 and not with PM? Vedal et al. state: "[t]hese findings provide no compelling evidence that short-term increases in relatively low concentrations of outdoor air pollutants have an adverse effect on individuals at risk of cardiac arrhythmias. The findings regarding SO_2 are difficult to interpret. They may be chance findings. Alternatively, given the very low concentrations of SO_2 that were present in Vancouver, SO_2 may have been serving as a surrogate measure of other environmental or meteorological factors."

A paper by Janes et al. (2007) is highly relevant to the issue of confounding. They concluded that the reported associations between PM and mortality may actually be attributable to inadequately controlled confounding. These authors used a novel approach to investigate confounding in air pollution studies. In the past two decades, there have been substantial decreases in air pollution across the United States, including fine $PM_{2.5}$ concentrations in parallel with decreases in death rates. It is not possible, however, to attribute the decline in death rates to a decline in pollution because of all the other changes in demographics, socio-economic factors and life-style that have also occurred over the same period of time. Janes et al. proposed an approach to addressing the issue of confounding. They argued that the association between national trends in fine PM and mortality "is likely to be confounded by slowly time-varying factors, such as changes in industrial activities and the economy, improving health care, and large scale weather events." However, at the local level these associations are less likely to be confounded. Therefore, they reasoned that if $PM_{2.5}$ is causally associated with mortality, then areas of the country that have seen large declines in $PM_{2.5}$ should also have larger declines in mortality than other areas of the country that experienced smaller declines in $PM_{2.5}$. To test the hypothesis that declines in $PM_{2.5}$ are causally associated with declines in mortality, they proposed a statistical approach that decomposed the association between $PM_{2.5}$ and mortality into a contribution at the national level and one at the local level. They analyzed the association between $PM_{2.5}$ and mortality in 113 U.S. counties over the three-year period 2000-2002, and concluded that "if the association at the national scale is set aside, there is little evidence of an association between 12-month exposure to $PM_{2.5}$ and mortality." This conclusion strongly

suggests that the reported associations between $PM_{2.5}$ and mortality are not causal but can be explained by confounding.

Thus, the Janes et al. study raises serious questions regarding the reported association between fine PM and mortality.

2. Long-Term Studies

While the time-series studies focus on the short-term health effects of air pollution, the long-term studies were conducted to assess the health impact of chronic exposure to air pollution on health. Because of the difficulties and expense of conducting long-term studies, only a small number have been done. As with time-series studies, the results have been mixed and inconsistent.

The time-series studies are based on day-to-day differences in ambient pollution levels while long-term studies use differences in pollutant levels between cities. Many more confounders must be controlled in long-term studies than in time-series studies. For example, smoking, which is not a confounder in time-series studies because an individual's status as a smoker or non-smoker does not vary day-to-day, is an important potential confounder in long-term studies. Because air pollution is measured on a city-wide level, any factor that is associated with mortality and varies from city to city, such as life-style or socio-economic factors, is a potential confounder in the long-term studies. Control of confounding by such factors can be extremely difficult. Adjustment for socio-economic factors is particularly difficult, and there is little assurance residual confounding bias can be eliminated. Thus all long-term studies need to be reviewed with particular caution.

The first long-term studies of particulate pollution and mortality were the Harvard Six Cities Study (Dockery et al., 1993) and the ACS study (Pope et al., 1995). While these studies attempted to control a number of individual-level confounders, such as cigarette smoking, they took no account of ecologic confounders, such as co-pollutants. As a consequence, the results of these studies were difficult to interpret, and a thorough re-analysis of the data was funded by HEI (Krewski et al., 2000). The re-analyses were conducted in two phases. In the first phase of replication of the original studies, the HEI investigators audited the data carefully and analyzed them using the methods used by the original investigators and essentially duplicated their results. Of much greater interest are the sensitivity analyses (phase 2) conducted by the HEI investigators on the ACS data (sensitivity analyses were not possible on the Six Cities data because of the small number of cities in the study). This phase explicitly considered a number of ecologic confounders, including co-pollutants.

a. The Harvard Six Cities Study Database

In the Harvard Six Cities Study, a random sample of over 8,000 adults was selected from six cities in Northeast and Midwest. Cox proportional hazards regression modeling, a commonly used statistical method, was used for analyses. Relative risks for mortality for residence in a particular city were estimated after adjustment for cigarette smoking, education and body mass index. In the city (Steubenville) with highest level of pollution as measured by levels of $PM_{2.5}$, the adjusted death rate was 26% higher than in the city (Portage) with the lowest pollution. If the

six cities are ranked in order of adjusted death rates from lowest to highest, and if this ordering is compared with the ordering imposed by various indices of air pollution, the agreement seems good, particularly if $PM_{2.5}$ is used as an index of air pollution. Thus, in this study, there appears to be good correlation between levels of $PM_{2.5}$ and death rates, after adjustment for some important confounders measured on the individual level.

However, in the Six Cities study, the RR for all-cause mortality of 1.26 reported for residents of Steubenville as compared to Portage is implausibly large. Using the risks associated with smoking estimated in the Six Cities study, a simple computation done by Moolgavkar (2010) shows that this RR is equivalent to increasing cigarette consumption among smokers by 25 pack-years. He concludes that the suggestion that ambient $PM_{2.5}$ contributes a risk as large as 25 pack-years of smoking is simply not biologically credible. Furthermore, the Six Cities study reports a RR of lung cancer mortality of 1.37 for residents of Steubenville as compared to Portage, which would make ambient $PM_{2.5}$ much more potent than direct emissions from coke ovens (Moolgavkar, 2005). This is also not biologically credible.

In addition, as Krewski et al. (2000) note in their re-analyses:

> The Six Cities Study, with its small number of cities and high degree of correlation among the air pollutants monitored, did not permit a clear distinction among the effects of gaseous and fine particle pollutants. Indeed, estimates of the relative risk of mortality from all causes were similar for exposure to fine particles, sulfate, sulfur dioxide, and nitrogen dioxide. Of the gaseous co-pollutants in the Six Cities Study, only ozone did not display an association with mortality.

In other words, SO_2 and nitrogen dioxide NO_2 are also correlated with mortality, but ozone is not. Moreover, the correlation may not reflect a causal association between air pollution and mortality at all. In fact, it was universally accepted at that time that SO_2 and NO_2 could not be causally associated with mortality in these studies. Rather, these associations may reflect uncontrolled confounding by covariates, such as measures of socio-economic status (e.g., income level, access to health care etc.). Thus, the correlation observed in the Six Cities Study does not even provide support for a causal association between exposure to air pollution generally and mortality, much less a causal association between $PM_{2.5}$ and mortality.

Laden et al. (2006) extended the Six Cities Study by eight years to 1998. They also used proportional hazards modeling to analyze the extended follow-up data and concluded that "[i]mproved overall mortality was associated with decreased mean $PM_{2.5}$..." However, Table 2 in the Laden paper shows that, in the period 1990-1998, mortality rates in the cities with high concentrations of $PM_{2.5}$ were no higher than the rates in cities with lower concentrations of $PM_{2.5}$.

If the relationship between $PM_{2.5}$ were linear without a threshold, then higher death rates should be seen in the cities with higher concentrations of $PM_{2.5}$ even in the follow-up period. This was not observed and, instead, the data suggest that below a concentration of about 22 $\mu g/m^3$, no $PM_{2.5}$ association with mortality is found. In addition, the average $PM_{2.5}$ concentration in period

2 in Watertown is approximately 2 $\mu g/m^3$ higher than in Portage, but the risk of death is lower in Watertown.

Other problems with this study are that death rates in the second period are being compared with death rates in the first among cohorts born 8 years earlier. Cohort effects are known to influence mortality strongly so that such a comparison is inappropriate. Also, the populations in the second period are 8 years older and relative risks are known to decrease with age (Villeneuve et al., 2002).

b. The ACS Database

The ACS Study (Pope et al., 1995) was a much larger study than the Six Cities Study as it involved 151 cities and more than 500,000 individuals. The design of the ACS Study was similar to the design of the Six Cities Study, and the study was undertaken specifically to test the major hypothesis raised by the Six Cities Study – that $PM_{2.5}$ was associated with mortality. With 151 cities in the data base, there was a real opportunity to control for confounders, particularly co-pollutants. However, the investigators did not do so and they did not explain why they failed to do so. The failure to consider confounding by co-pollutants is a major deficiency of this study and casts doubt on the reported association between fine PM and mortality.

In response to the shortcomings of Pope et al., Krewski et al. (2000) explicitly considered a number of confounders, including co-pollutants. Some important findings of their analyses were: (1) substantial attenuation of the $PM_{2.5}$/mortality association occurred when SO_2 was considered in a two-pollutant model; (2) attenuation of the PM effect also occurred when spatial correlation was considered, which suggests that the reported association with $PM_{2.5}$ is spurious; and (3) modification of the $PM_{2.5}$ association by level of education (i.e. the association was observed only in the subpopulation with a high school education or less). This last finding suggests that important socio-economic confounders have not been identified or controlled for, and cast doubt on the validity of a causal $PM_{2.5}$/mortality relationship.

In the second update of the ACS study, Pope et al. (2004) analyzed the association between $PM_{2.5}$ and cardiovascular mortality and they concluded that "[a]lthough smoking is a much larger risk factor for cardiovascular disease mortality, exposure to fine PM imposed additional effects that seem to be at least additive if not synergistic with smoking." They also concluded that the association of $PM_{2.5}$ with respiratory mortality was weak. In fact, however, these authors found a statistically significant negative (protective) association between exposure to $PM_{2.5}$ and respiratory mortality. This finding does not seem biologically plausible and casts doubt on their other reported findings. In addition, their reported finding of a possible synergistic action between $PM_{2.5}$ and smoking with mortality conflicts with the results of an earlier paper (Pope et al., 2002). In that paper, they reported that fine PM associated mortality risks were lower among smokers than among non-smokers. Finally, the authors again did not consider any pollutants other than $PM_{2.5}$ in their analyses.

Another study (Jerrett et al., 2005) that used the ACS database also failed to adequately account for important confounders. Jerrett et al. examined data for about 23,000 subjects in the Los Angeles Basin from the ACS cohort for the period 1982–2000, with more than 5,000 deaths.

Pollution exposures were estimated from 23 $PM_{2.5}$ and 42 ozone monitors. After controlling for 44 individual risk factors for mortality (e.g., smoking), they found a significantly increased risk of mortality associated with $PM_{2.5}$ for all-cause, ischemic heart disease, and lung cancer. The only other pollutant considered was ozone, and the authors found that the $PM_{2.5}$ results were not affected by including ozone or an adjustment for expressway exposure. The authors also found that the magnitude of fine $PM_{2.5}$ effects are about three times as large as those found in earlier studies. They implied that this was due to better exposure estimates obtained by interpolation of the pollution data to a finer scale and suggested that the chronic health effects associated with within-city gradients in exposure to $PM_{2.5}$ may be even larger than previously reported across metropolitan areas.

However, when covariates related to socioeconomic status were included in the analyses, the associations of $PM_{2.5}$ with total, ischemic heart disease, and lung cancer mortality were substantially attenuated and became either insignificant or only borderline significant. Despite the finding that SO_2 was associated with mortality in Krewski et al. (2000), Jerrett et al did not consider co-pollutants other than ozone. This oversight is significant because in time-series analyses in Los Angeles, both CO and SO_2 have been found to be associated with mortality even though concentrations of SO_2 are low in LA (Moolgavkar, 2000, 2003a, b). Finally, the RR for lung cancer in this study (1.44 without covariates) is much higher than that reported in any of the previous analyses of the ACS cohort which, as discussed above, is much too high to be biologically plausible. Unfortunately, the paper does not present the relative risks associated with strong risk factors, such as cigarette smoking, estimated in this study. In epidemiology studies, the estimated risks from such factors can be used as reality checks of whether the analyses yield reasonable estimates of well-studied risk factors.

The final analyses of the ACS cohort was by Krewski et al. (2009). They increased the follow-up period of the ACS cohort to 18 years and conducted a finer-scale spatial analyses in Los Angeles and New York. For the overall analyses, the study reports a RR for total mortality of 1.03 associated with a 10 $\mu g/m^3$ increase in $PM_{2.5}$, which is much smaller than the RR found in other analyses of the ACS cohort.

However, this study, like others, fails to account for important confounders. These analyses again show, as in the 2000 re-analyses of the ACS study, that even with 18 additional years of follow-up, other pollutants still show a statistically significant relationship with mortality in this data set. Despite this fact and the finding in the 2000 re-analyses that other pollutants were much more strongly associated with mortality than either $PM_{2.5}$ or sulfates, the investigators of this study did not report the results of any two-pollutant models.

The investigators also conducted finer-scale spatial analyses in Los Angeles and New York. The purpose was to reconstruct $PM_{2.5}$ levels on a local basis with a more accurate exposure assessment. The observed associations between $PM_{2.5}$ and total mortality and ischemic heart disease (IHD) mortality were strong and significant in Los Angeles. However, in New York there was no association between $PM_{2.5}$ and total mortality but there was a statistically significant association between $PM_{2.5}$ and IHD mortality. Since IHD mortality comprises a high fraction of total mortality, especially at the older ages, these findings do not seem plausible, particularly in view of the large $PM_{2.5}$ coefficient reported for IHD mortality.

As mentioned above, New York was one of the two metropolitan areas in the U.S. where a statistically significant association was reported between PM and all cause mortality in NMMAPS. Yet, in this fine-scale spatial analysis in New York, no association was detected between $PM_{2.5}$ and all-cause mortality. These disparate results point, once again, to the general inconsistency in the air pollution literature.

c. Other Long-Term Studies

In addition to the studies using the Six Cities and ACS cohorts, a few other notable long-term studies have appeared (Lipfert et al., 2000, 2006a, b, 2009; Enstrom, 2005). These latter studies arrive at conclusions that are fundamentally at odds with those cited above because of the additional variables included in their analyses.

B. Exposure Uncertainty

An ideal epidemiology study would have information on both exposure and disease outcome on each individual in the study. Most air pollution epidemiology studies do not even come close to having this type of information. Concentrations of pollutants in the air are obtained from air monitoring sites at locations scattered throughout the United States. These monitors have generally been sited to determine compliance with the NAAQS. In epidemiology studies, the concentrations measured at these sampling stations are assumed to be representative of actual exposures received by individuals living near them. This assumption is false. Pollutant concentrations can and do exhibit significant spatial (horizontal as well as vertical) and temporal gradients. Actual exposure is determined by where people live, where they work, the time they spend indoors and outdoors, and the myriad habits of daily life. In the 2004 PM CD, EPA points out that people spend about 90% of their time indoors and only about 6% of their time outdoors where they are directly exposed to ambient PM. The general population spends between 50 and 60% of the time at their place of residence. For the frail population in hospital and nursing homes, the time spent indoors approaches 100%. Thus, the exposure of individuals varies greatly, and central monitors do not capture this variation, creating significant uncertainty in exposure estimates in epidemiological studies.
.
The problem of exposure uncertainty is worse when direct measurements of the pollutant of interest are not available from the central monitors. For example, population-based monitors were employed in the original Six Cities study. However, the monitors were in operation only for 1979-1987. For the follow-up of the Six Cities study (Laden et al., 2006), no direct measurements were available. Thus, over the period 1987-1998, $PM_{2.5}$ concentrations were estimated using regression equations, not actual data. The extra variability introduced by this procedure was never addressed in the Laden et al. analyses of the Six Cities data.

Two main study designs have been used in air pollution epidemiology: one for the investigation of short-term effects and the other for the investigation of long-term associations. Time-series studies are designed to investigate the short-term effects of air pollution on human health. In a time-series study, fluctuations in concentrations of specific air pollutants from day to day are statistically related to the total number of health events in a population, such as hospital

admissions and deaths on subsequent days. Time-series studies of air pollution suffer from another significant source of uncertainty because they rely on summary measures of health in a population, such as the total number of hospital admissions or of deaths on any given day. *There is no assurance that **any** of the individuals who died or became ill were actually exposed to the highest levels of pollution.*

Long-term studies designed to observe the chronic effects of air pollution on health rely not on daily air pollution measurement, but annual averages and compare the averages between different cities. Only a relatively small number of studies have investigated the long-term association of air pollution and mortality/morbidity. These studies obtain data on some potential confounders, such as cigarette smoking, and the vital statistics are available on each individual in the study. However, exposure to air pollution is available only from the closest central monitors. *Thus, there is no way of assessing the actual exposures of any individuals who died over the course of the study.*

On page 2-25 of the PM ISA, EPA states: "Overall, the limited evidence from the studies evaluated supports the use of a no-threshold, log-linear model, which is consistent with the observations made in studies that examined the PM-mortality relationship." However, as shown below, measurement error associated with exposure likely obscures the evidence for a threshold.

The shape of the concentration-response function and the existence of a threshold were major considerations during the review of the 2004 PM Criteria Document and the development of the risk assessment included in the 2005 PM Staff Paper. Although early drafts of the CD indicated that the PM studies generally show linear concentration-response associations, the final CD concludes on page 9-44 that "In summary, the available evidence does not either support or refute the existence of thresholds for the effects of PM on mortality across the range of concentrations in the studies." They also note on page 8-320: "the available information does not allow for a clear choice of 'threshold' or 'no threshold' over the other." This view is consistent with points made by the Special Panel of the HEI Review Committee (Special Panel of the Health Review Committee, 2004) that raised several cautions in interpreting the NMMAPS concentration-response results. They point out that measurement error could obscure any threshold that might exist, that city-specific concentration-response curves exhibited a variety of shapes, and that the use of Akaike Information Criterion may not be an appropriate criterion for choosing between models. The HEI Panel cautioned that lack of evidence against a linear model should not be confused with evidence in favor of it. In addition, Rhomberg et al. (2011a) have recently shown, as others have previously shown, that measurement error can give a false linear result. Thus, the epidemiological studies cannot inform us as to whether there is or is not a biologic gradient for ambient PM at low concentrations or whether there is or is not a threshold.

The toxicological studies of PM components that have been used to set chemical–specific standards demonstrate both threshold behavior and the presence of effects that not only become less common with progressively lower doses, but they also become less severe. The existence of a substantial threshold for the first physiological effects in controlled studies is not consistent with the assumption that the more severe effects suggested by some epidemiological studies have no threshold. Such assumptions are not consistent with either the general principles of

toxicology or the specific findings of PM toxicological studies. Rhomberg et al. (2011b) discusses these issues in detail:

> The no-threshold proposal for noncancer toxicity is at variance with decades of experience in observing exposure-response relationships in pharmacology and toxicology, both within and below the usual experimental range for environmental chemicals.

They further note:

> The no-threshold idea is also belied by our experience with medicines, poisons, foodstuffs, and many other kinds of exposure to agents that can have toxic effects if experienced in excess. With the possible exception of allergic reactions, within the range of low exposures, we do not observe slightly increased exposures to such agents somewhat increasing the probability that we will suffer the full effect of a toxic dose. In therapeutics, a small fraction of the therapeutic dose will not necessarily produce a moderate or full response in a diminished fraction of the treated population. It is only when the critical concentration is sustained at the site of action for the necessary period of time that an effect will be elicited. The experience of exposure thresholds for biological effects, including adverse effects, pervades daily life.

They also argue that the no-threshold proposal is at variance with basic tenets of homeostasis—the robust nature of living systems.

In summary, the shape of the concentration-response is not known and epidemiology studies cannot be used to identify threshold because of exposure uncertainty. Consequently EPA's extrapolations of risk at low PM concentrations are inappropriate.

C. Model Selection Bias

In epidemiology, statistical models are used to relate a health outcome to various factors that may contribute to the occurrence of that health outcome. Selecting an appropriate statistical model for epidemiology analyses of air pollution data is an extremely important process that can affect the outcome of the study in a very significant way. It can make the difference between finding a positive association, a negative association or no association. It involves making a number of choices which include:

- How is confounding by weather to be controlled? That is, what functional form should be assumed for the effects of weather variables, such as temperature and relative humidity?
- What weather variables should be used?
- What co-pollutants should be included and what averaging time should be used?
- What temporal effects need to be controlled and to what degree?
- What lag structure should be assumed? That is, how many days after exposure to a pollutant should one expect to see an effect on health?

There is little biological knowledge to inform these choices that must be made.

In a commentary on the challenges of air pollution epidemiology, Lumley and Sheppard (2003) point out:

> Estimation of very weak associations in the presence of measurement error and strong confounding is inherently challenging. In this situation, prudent epidemiologists should recognize that residual bias[5] can dominate their results. Because the possible mechanisms of action and their latencies are uncertain, *the biologically correct models are unknown*. This model selection problem is exacerbated by the common practice of screening multiple analyses and *then selectively reporting only a few important results* (emphasis added).

Many others have made similar comments regarding the critical importance of model choice, particularly when effect estimates are small, which they are in air pollution epidemiology studies. For example, in comments on a draft PM CD submitted to the EPA, Smith, et al. (2001), state:

> 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). As has been demonstrated in Lumley and Sheppard (2000), the effect of choosing lags for PM in this fashion has a bias which is of the same order of magnitude as the relative risk being estimated (emphasis added).

Koop and Tole have been especially outspoken in their concerns over model selection bias. Koop and Tole (2004) state:

> The main empirical finding of [our] paper is that standard deviations for air pollution-mortality impacts become very large when model uncertainty is incorporated into the analysis. *Indeed they become so large as to question the plausibility of the previously measured links between air pollution and mortality* (emphasis added).

The main conclusion from their paper was that when model uncertainty was considered, there was little evidence of a PM association with mortality.

A single event was responsible for raising the appreciation of the model selection bias issue more than any one single paper. That event occurred in May of 2002. Most time series studies of air pollution had used the Generalized Additive Models (GAM) for analyses of data. The most widely used software for fitting these models is a statistical package called S-plus. In May of

[5] Residual bias is the bias that may remain by chance after all known sources of bias have been controlled.

2002, the NMMAPS investigators discovered that the implementation of GAM in S-plus was flawed and could yield misleading results (Health Effects Institute, 2003b). EPA, which was in the process of preparing a revised PM CD, halted work on the CD and asked investigators to re-analyze a number of studies that EPA had identified as key studies. These re-analyses were carried out under the auspices of the Health Effects Institute and published in 2003 (HEI, 2003b) with commentaries by the expert panel (Special Panel of the Health Review Committee, 2003) convened by HEI to serve as a peer review panel for the revised analyses.

The revised analyses necessitated by the S-plus problems clearly indicate that methods used for controlling temporal trends and weather can have profound effects on the results of time-series analyses of air pollution data, as the HEI expert panel noted (HEI, 2003b; pages 267, 269). Moreover, there appears to be no objective statistical test to determine whether these factors have been adequately controlled in any analysis. The HEI Expert Panel for the re-analyses stated on page 267:

> Ritov and Bickel (1990) have shown, however, that for any continuous variable, no strictly data-based (*i.e.*, statistical) method can exist by which to choose a sufficient number of degrees of freedom to insure that the amount of residual confounding due to that variable is small. This means that no matter what statistical method one uses to select the degrees of freedom, *it is always logically possible that even if the true effect of pollution is null, the estimated effect is far from null due to confounding bias* (emphasis added).

In other words, even if the true effect of pollution is zero, the estimated effect may be positive because it is impossible to control temporal trends or weather without accurate information from external sources that do not exist. The HEI expert panel (HEI, 2003; page 269) comments further*, "Neither the appropriate degree of control for time, nor the appropriate specification of the effects of weather, has been determined for time-series analyses"* (emphasis added).

The HEI Special Review Committee summarized the overall impact of correcting the GAM criteria on the studies reanalyzed:

- While the number of studies showing an association of PM with mortality was slightly smaller, the PM association persisted in the majority of studies.
- In some of the large number of studies in which the PM association persisted, the estimates of PM effect were substantially smaller.
- In the few studies in which investigators performed further sensitivity analyses, some showed marked sensitivity of the PM effect estimate to the degree of smoothing and/or the specification of weather.

One of the re-analysis participants tested the impact of model selection by running over a thousand possible models. Ito (2003) carried out a systematic re-analysis of the air pollution associations in the Detroit area and re-analyzed the 1220 separate air pollution mortality and morbidity associations that were included in the original Lippmann et al. (2000) study of Detroit.

As shown in Figure 1 of his report, there was a wide range of negative and positive risks in Detroit when all pollutants, lags, and endpoints were considered. Ito showed that the wide range of associations occurred for each pollutant. Although the focus in the original Lippmann et al. study, as it is in almost all the published literature, was on the positive associations, Ito's plot showed that there are many negative associations in the data. Although there may be somewhat more positive associations than negative associations, there is so much variability in the risk estimates, that identifying which positive associations may be real health effects and which are not appears beyond the capability of current methods. Moreover, in the Ito re-analysis, the overall pattern for each pollutant is similar so that one pollutant or one PM indicator is not implicated over any of the others.

A final paper on model selection bias that deserves attention is another contribution from Koop and Tole. Koop et al. (2010) underscores many of the issues raised in the preceding paragraphs and adds additional insights as to the reasons why the real relationships between health effects and air pollution at relevant exposures are small and insignificant. In this study, the authors conduct a comprehensive analysis of air pollution morbidity relationships for eleven Canadian cities over a long record from 1974 to 1994. As a result, they have a unique data set that allowed the examination of both spatial and temporal variations. In addition to including the five criteria pollutants, CO, PM, SO_2, NO_2 and O_3, they also controlled for socioeconomic factors, smoking and meteorology. Much shorter subsets of this data set have been analyzed without the socioeconomic and smoking variables by a number of research groups to demonstrate significant relationships with a number of health outcomes and individual pollutants. The long data set enabled the present investigators to explore the impact of significantly lower air pollution concentrations at the end of the data set compared to the beginning. Koop et al. also employed the two major methods used to formulate the statistical models in time-series studies, model selection by the use of some statistical criteria and Bayesian Model Averaging (BMA), to address the all-important issue of model selection uncertainty.

As Koop et al. noted for air pollution/mortality or morbidity epidemiology results in general, the results are conflicted. In other words, the results range from positive to negative and from significant to insignificant for all pollutants and for all health endpoints. Koop et al. state:

> One of the reasons for this profusion of apparently contradictory results is model uncertainty. With very few exceptions (e.g. Clyde, 2000; Clyde and DeSimone-Sasinowska, 1997 and Koop and Tole, 2004, 2006), previous studies on air pollution-health effects have used model selection methods, i.e. choosing one or a few regression specifications and reporting point estimates and their associated variances conditional on that being the true model. However, the estimation exercise is inherently opportunistic. Many plausible covariates could be included, but the choice is not dictated by theory so much as by data availability. Hence there is not only uncertainty about regression slope coefficients conditional on the model selection, but about the model specification itself.

Compounding the issue of selecting the true model is the large number of potential explanatory variables and possible forms that will influence the model results. As Koop et al. articulate:

> However, the number of potential confounding variables implies that a huge number of models could be used to explain health effects. The number of potential models is on the order of 2^k where k is the number of potential explanatory variables, including lags. Since results can be sensitive to the particular regression specification, and since the number of potential models is so large, model uncertainty has been shown to be an important issue in this literature (Clyde, 2000; Koop and Tole, 2004).

To address the model uncertainties, Koop and Tole use BMA. This method includes information from every potential model. The BMA results are weighted averages of the estimates from each model. The weights are proportional to the support the data give each model.

The results of the BMA analyses show that the health outcomes are explained by the smoking and the socioeconomic variables and that none of the air pollutants showed a statistically positive relationship with health. In fact most pollutant relationships were slightly negative, but not robust. With this particular data set the BMA results were largely similar (except NO_2 showed an effect in a single model) to the results obtained by selecting a single model. This is in contrast to their earlier results (Koop and Tole, 2004) for Toronto which found many relationships when a single model was used. In the earlier paper, a shorter data record was used and the smoking and socioeconomic variables were not included. This may explain the differences and underscores the importance of including these variables in a longer time-series in these types of studies.

In summary, this study demonstrates the importance of: 1) incorporating smoking and socioeconomic variable into the models, 2) using a longer time series that has significantly different pollutant concentrations at the beginning and end of the study, 3) using the BMA approach which minimizes model selection uncertainties and finds insignificant health impacts. This suggests that the epidemiological evidence relied on by EPA is scientifically unsound and should not be used as a reason to drive the NAAQS lower and lower.

References

77 Fed. Reg. 38890-39055 (June 29, 2012).

77 Fed. Reg. 38900 (June 29, 2012).

77 Fed. Reg. 38943 (June 29, 2012).

Anderson, H., R. Atkinson, J. Peacock, M. Sweeting, and L. Marston. 2005. Ambient particulate matter and health effects: publication bias in studies of short-term associations, *Epidemiology*, 16:155-163.

Clyde, M. 2000. Model uncertainty and health effect studies for particulate matter. *Environmetrics* 11:745–764.

Clyde, M. and H. DeSimone-Sasinowska, H., 1997. Accounting for Model Uncertainty in Poisson Regression Models: Particulate Matter and Mortality in Birmingham, Alabama. Institute of Statistics and Decisions Sciences, Duke University Discussion Paper 97-06.

Dockery, D.W., C.A. Pope, X. Xu, J.D. Spengler, J.H. Ware, M.E. Fay, B.G. Ferris and F.E.Speizer. 1993. An association between air pollution and mortality in six U.S. cities. *N. Engl. J. Med.*, 329:1753-1759.

Dominici F., A. McDermott, M. Daniels, S.L. Zeger and J.M. Samet.. 2003, Revised analysis of the National Morbidity, Mortality, and Air Pollution Study, Part II. In: *Revised Analyses of Time-Series Studies of Air Pollution and Health,* HEI Special Report, pp. 5-24.

Dominici F., D. Peng, M. Bell, A. Pham, A. McDermott, S.L. Zeger and J.M. Samet. (2006). Particles, air pollution and hospital admissions for cardiovascular and respiratory diseases. *J. American Medical Association*, 295:1127-1134.

Dominici, F., R.D. Peng, S.L. Zeger, R.H. White and J.M. Samet. 2007. Particulate air pollution and mortality in the United States: did the risks change from 1987 to 2000? *Am. J. Epidemiol*, 166: 880-888.

Enstrom, J.E. 2005. Fine particulate air pollution and total mortality among elderly Californians, 1973-2002. *Inhalation Toxicol*., 17:803-816.

Franklin, M., A. Zeka and J. Schwartz. 2007. Association between PM2.5 and all-cause and specific-cause mortality in 27 US communities. *J. Expo. Sci. Environ. Epidemiol*., 17: 279-287.

Godleski, J.J., A.C. Rohr, B.A. Coull, C-M.Kang, E.A. Diaz and P. Koutrakis. 2011. Toxicological evaluation of realistic emission source aerosols (TERESA): summary and conclusions. *Inhalation Toxicology*,23(S2):95-103.

Goodman, S. 2005. The Methodologic ozone effect. *Epidemiology,* 16: 430-435.

Health Effects Institute. 2003a. *Health Effects of Acute Exposures to Air Pollution.* HEI Publication # 112. 67pp.

Health Effects Institute. 2003b. *Revised Analyses of Time-Series Studies of Air Pollution and Health.* HEI Special Report. 291pp.

Ito, K. 2003. Associations of particulate matter components with daily mortality and morbidity in Detroit, MI. In: *Revised Analyses of Time-Series Studies of Air Pollution and Health,* HEI Special Report., pp. 143-156.

Janes, H., F. Dominici and S.L. Zeger. 2007. Trends in air pollution and mortality – An approach to the assessment of unmeasured confounding. *Epidemiology*, 18:416-423.

Jerrett, M., R.T. Burnett, R. Ma, C.A. Pope, D. Krewski, K.B. Newbold, G. Thurston, Y. Shi, N. Finkelstein, E.E. Calle and M.J. Thun. 2005. Spatial analysis of air pollution and mortality in Los Angeles. *Epidemiology*, 16: 727-736.

Kang, C-M., T. Gupta, P.A. Ruiz, J.M. Wolfson, S.T. Ferguson, J.E., Lawrence, A.C. Rohr, J. Godleski and P. Koutrakis. 2010. Aged particles derived from emissions of coal-fired power plants: The TERESA field results. *Inhalation Toxicology*, DOI:10.3109/08958371003728040.

Katsouyanni K. and J. Samet. (2009*). Air Pollution and Health: A European and North American Approach (APHENA)*, Health Effects Institute Report 142, Cambridge, MA.

Koop G. and L. Tole. 2004. Measuring the health effects of air pollution: to what extent can we really say that people are dying from bad air? *J. Environ Econ and Management*, 47:30-54.

Koop, G. and L.Tole. 2006. An Investigation of thresholds in air pollution mortality effects. *Environmental Modelling & Software*. 21:1662–1673.

Koop, G., R. McKitrick, and L. Tole, 2010. Air pollution, economic activity and respiratory illness: Evidence from Canadian cities, 1974-1994. *Environ. Model. Softw*. 25:873-885.

Krewski, D., R.T. Burnett, M.S. Goldberg, K. Hoover, J. Siemiatycki, M. Jerrett, M. Ambramowicz, and W. H. White. 2000. *Reanalysis of the Harvard Six Cities Study and the American Cancer Study of Particulate Air Pollution and Mortality,* Health Effects Institute Special Report, Cambridge, MA.

Krewski D.et al. 2009. *Extended follow-up and spatial analysis of the American Cancer Society study linking particulate air pollution and mortality*. Research Report 140. Health Effects Institute, Boston, MA.

Laden, F., J. Schwartz, F.E. Speizer and D.W. Dockery. 2006. Reduction in fine particulate air pollution and mortality: extended follow-up of the Harvard Six Cities study. *Am. J. Respir Crit. Care Med.,* 173: 667-672.

Lipfert, F.W., H.M. Perry, J.P. Miller, J.D. Baty, R.E. Wyzga and S.E. Carmody. 2000. The Washington University-EPRI veterans' cohort mortality study: preliminary results. In: Grant, L. D.(ed.) PM 2000: particulate matter and health. *Inhalation Toxicol*. 12:41-73.

Lipfert, F.W., R.E. Wyzga, J.D. Baty and J.P. Miller. 2006a. Traffic density as a surrogate measure of environmental exposures in studies of air pollution health effects: long-term mortality in a cohort of U.S. veterans. *Atmospheric Environment*, 40:154-169.

Lipfert, F.W., J.D. Baty, J.P. Miller and R.E. Wyzga. 2006b. PM constituents and related air quality variables as predictors of survival in a cohort of U.S. military veterans. *Inhalation Toxicol.,* 18:645-667.

Lipfert, F.W., R.E. Wyzga, J.D. Baty and P.J. Miller. 2009. Air pollution and survival within the Washington University-EPRI Veterans cohort: Risks based on modeled estimates of ambient levels of hazardous and criteria air pollutants. *J. Air & Waste Mgt. Assoc*., 60:473-489.

Lippmann M., K. Ito, A. Nádas and R.T. Burnett RT. 2000. *Association of Particulate Matter Components with Daily Mortality and Morbidity in Urban Populations.* Research Report 95, Health Effects Institute, Cambridge MA.

Lumley T. and L.Sheppard. 2000. Assessing seasonal confounding and model selection bias in air pollution epidemiology using positive and negative control analysis. *Environmetrics*, 11:705-717.

Lumley T. and L.Sheppard. 2003. Time series analyses of air pollution and health: straining at gnats and swallowing camels? *Epidemiology*, 14:13-14.

Miller, K.A., D.S. Siscovick, L.Sheppard, K. Shepherd, J.H. Sullivan, G.L. Anderson and J.D. Kaufman. 2007. Long-term exposure to air pollution and incidence of cardiovascular events in women. *N. Engl. J. Med.*, 356: 447-458.

Moolgavkar, S.H. 2000. Air pollution and mortality in three U.S. counties. *Environ. Health Perspect.*, 108:777-784.

Moolgavkar, S.H. 2003a. Air pollution and daily deaths and hospital admissions in Los Angeles and Cook counties. In: *Revised analyses of time-series studies of air pollution and health. Special report.* Health Effects Institute, Boston, MA. pp. 183-198.

Moolgavkar, S.H. 2003b. Air pollution and daily mortality in two U.S. counties: season-specific analyses and exposure-response relationships. *Inhalation Toxicology*, 15:877-907.

Moolgavkar, S.H. 2005. A review and critique of the EPA's rational for a fine particulate standard. *Regulatory Toxicology & Pharmacology*, 42:123-144.

Moolgavkar, S.H. (2010). Personal Communication.

National Research Council. 2004. *Research Priorities for Airborne Particulate Matter. IV. Continuing Research Progress.*Report of the National Research Council of the National Academies.

Ostro B., R. Broadwin, S. Green, W.-Y. Feng and M. Lipsett. 2006. Fine particulate air pollution and mortality in nine California counties: results from CALFINE. *Environ. Health Perspect*., 114: 29-33.

Peng, R. D., F.Dominici, R. Pastor-Barriuso, S.L. Zeger and J.M. Samet. (2005). Seasonal analyses of air pollution and mortality in 100 U. S. Cities. *Am. J. Epidemiol.*, 161:585-594.

Pope, C.A., M.J. Thun, M.M. Namboodiri, D.W. Dockery, J.S. Evans, F.E. Speizer and C.W. Heath. 1995. Particulate air pollution as a predictor of mortality in a prospective study of

U.S. adults. *Am. J. Respir. Crit. Care Med*., 151: 669-674.

Pope, C.A., R.T. Burnett, M.J. Thun, E.E. Calle, D. Krewski, K.Ito and G.D. Thurston. 2002. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *J. Am. Med. Assoc*., 287:1132-1141.

Pope, C.A., R.T. Burnett, G.D. Thurston, M.J. Thun, E.E. Calle, D. Krewski and J.J. Godleski. 2004. Cardiovascular mortality and long-term exposure to particulate air pollution. *Circulation,* 109:71-77.

Pope, C.A. and R.T. Burnett. 2007. Confounding in air pollution epidemiology: the broader context. *Epidemiology*, 18(4):424-426.

Reiss, R., E.L. Anderson, C.E. Cross, G. Hidy, D. Hoel, R. McClellan and S. Moolgavkar. 2007. Evidence of health impacts of sulfate- and nitrate-containing particles in ambient air. *Inhalation Toxicology,* 19:419-449.

Rhomberg, L.R., J. K. Chandalia, C. M. Long, and J. E. Goodman. 2011a). Measurement error in environmental epidemiology and the shape of exposure-response curves. *Critical Reviews in Toxicology,* 41:651-671.

Rhomberg, L.R., J. Goodman, L.Haber, M. Dourson, M. Andersen, J. Klaunig, B. Meek, P. Price, R. McClellan and S. Cohen. 2011b). Linear low-dose extrapolation for noncancer health effects is the exception, not the rule. *Crit.Rev.Toxicol*., 41:1-19.

Ritov Y. and P. Bickel. 1990. Achieving information bounds in non- and semi-parametric models. *Ann. Stat*., 18:925–938.

Rochester Conference Report. 2007. *Critical Considerations in Evaluating Scientific Evidence of Health Effects of Ambient Ozone,* University of Rochester School of Medicine.

Schlesinger, R.B. and F. Cassee. 2003. Atmospheric secondary inorganic particulate matter: the toxicological perspective as a basis for health effects risk assessment. *Inhalation Toxicology,* 15:197-235.

Smith R., P. Guttorp, L. Sheppard, T Lumley and N.Ishikawa. 2001. *Comments on the Criteria Document for Particulate Matter Air Pollution*, NRSCE Technical Report Series #66, July 25, 2001. Available:http://www.nrcse.washington.edu/research/reports.html (as of 7-28-2012).

Snipes, M. B., A.C. James and A.M. Jarabek. 1997. The 1994 ICRP66 human respiratory tract dosimetry model as a tool for predicting lung burdens from exposures to environmental aerosols. *Appl. Occup. Environ. Hyg.*12: 547-554.

Special Panel of the Health Review Committee. 2003. Commentary on Revised Analyses of selected studies. In: *Revised Analyses of Time-Series Studies of Air Pollution and Health,* HEI Special Report, pp. 255-291.

Special Panel of the Health Review Committee. 2004. Commentary. In: *The National Morbidity, Mortality, and Air Pollution Study Part III: Concentration-Response Curves and Threshold for the 20 Largest US Cities,* HEI Report 94, Part III, pp. 23-30.

U.S. EPA. 1996. *Review of the National Ambient Air Quality Standards for Particulate Matter: Policy Assessment of Scientific and Technical Information OAQPS Staff Paper.* EPA-452\R-96-013, Research Triangle Park, NC.

U.S. EPA. 2004. *Air Quality Criteria for Particulate Matter, Vol. II.* EPA-600\P-99-002bF, Research Triangle Park, NC.

U.S. EPA. 2005. *Review of the National Ambient Air Quality Standards for Particulate Matter: Policy Assessment of Scientific and Technical Information OAQPS Staff Paper.* EPA-452\R-05-005a, Research Triangle Park, NC.

U. S. EPA. 2008a. *Second External Review Draft of Integrated Science Assessment for Oxides of Nitrogen-Health Criteria*, EPA-600/R-07/093aB, Research Triangle Park, NC.

U. S. EPA. 2008b. *Integrated Science Assessment for Oxides of Sulfur-Health Criteria*, EPA-600/R-07/047F, Research Triangle Park, NC.

U.S. EPA. 2009. *Integrated Science Assessment for Particulate Matter.* EPA-600\R-08/139F, Research Triangle Park, NC.

U.S. EPA. 2010. *Quantitative Risk Assessment for Particulate Matter.* EPA-452/R-10-005, Research Triangle Park, NC.

U.S. EPA. 2011. *Policy Assessment for the Review of the Particulate Matter National Ambient Air Quality Standards.* EPA-452\R-11-003, Research Triangle Park, NC

U.S. EPA. 2012. Particulate Matter (PM) Standards - Table of Historical PM NAAQS. http://www.epa.gov/ttn/naaqs/standards/pm/s_pm_history.html (accessed July 26, 2012).

Valberg, P.A. 2004. Is PM more toxic than the sum of its parts? Risk-Assessment toxicity factors vs. PM-mortality "effect functions", *Inhalation Toxicology,* 16(suppl.1):19–29.

Vedal S., K. Rich, M. Brauer, R. White and J.Petkau. 2004. Air pollution and cardiac arrhythmias in patients with implantable cardioverter defibrillators. *Inhalation Toxicology* 16:353-362.

Villeneuve, P.J., M.S. Goldberg, D. Krewski, R.T. Burnett and Y. Chen. 2002. Fine particulate air pollution and all-cause mortality within the Harvard Six-Cities Study: Variations in risk

by period of exposure. *Ann. Epidemiol.*, 12:568-576.

Vostal, J. 2000. Statistical associations between ambient particulate matter and daily morbidity and mortality: Can we identify mechanisms responsible for these health effects? *Proceedings of the Air & Waste Management Association's 93rd Annual Conference & Exhibition*, Publ. VIP 97, Air & Waste Management Association, Pittsburgh, PA.

Winter-Sorkina, R. de and F.R. Cassee. 2002. *From concentration to dose: factors influencing airborne particulate matter deposition in humans and rats*. Bilthoven, The Netherlands National Institute of Public Health and the Environment (RIVM), report no. 650010031/2002. Available: http://www.rivm.nl/bibliotheek/rapporten/650010031.html (13 June 2003).

Zeger S., F. Dominici, A. McDermott and J. Samet. 2008. Mortality in the Medicare population and chronic exposure to fine particulate air pollution in urban centers (2000-2005). *Environ. Health Perspect.*, 116:1614-1619.

Appendix A

Figures and Table

Summary of NMMAPS Individual City Results

Figure A1
NMMAPS Maximum Likelihood Estimates and 95% Confidence Intervals of the Percentage Increase in Total Mortality from Nonexternal Causes per 10 ppb Increase in Ozone Concentration for Each Location
Lag 1 Day

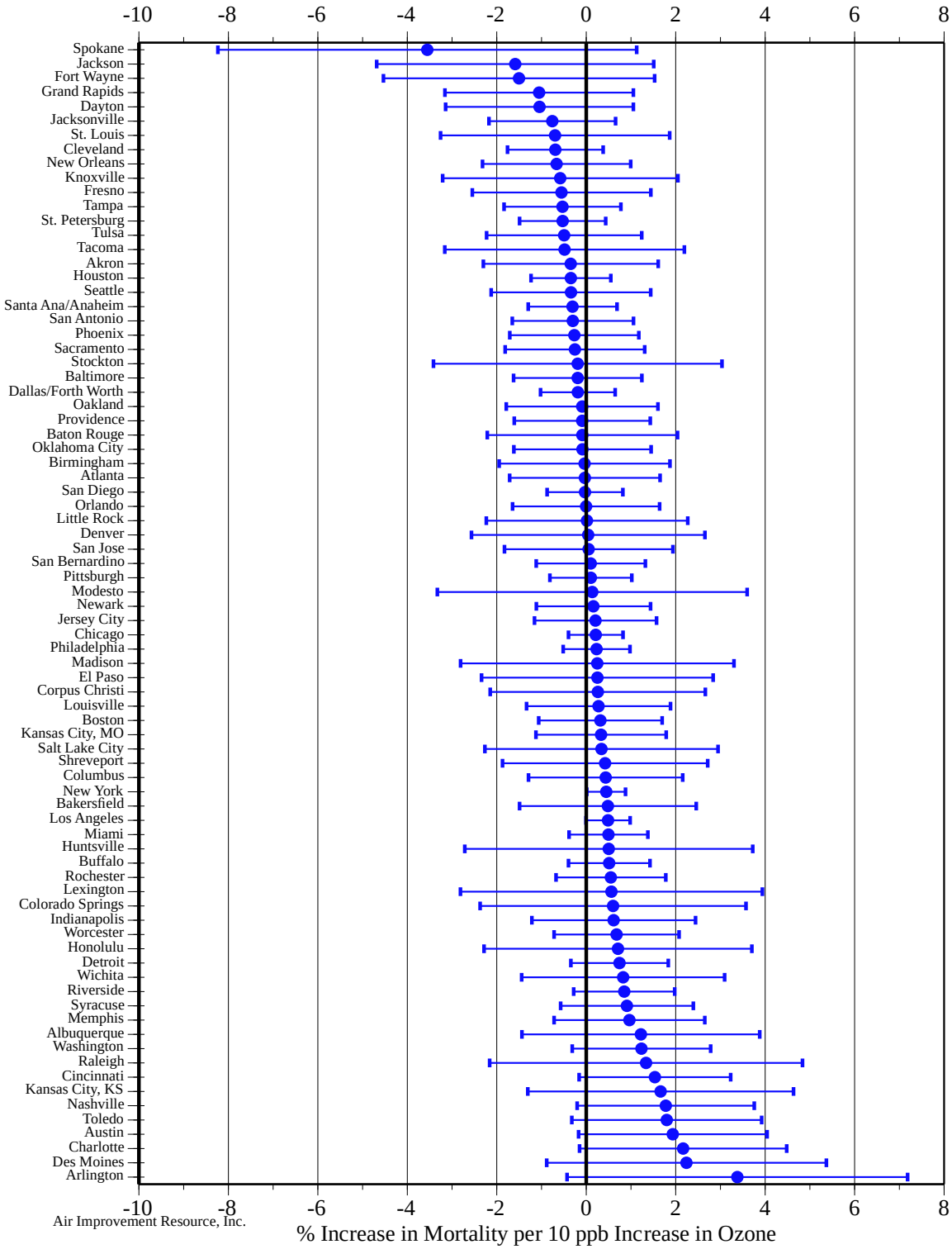
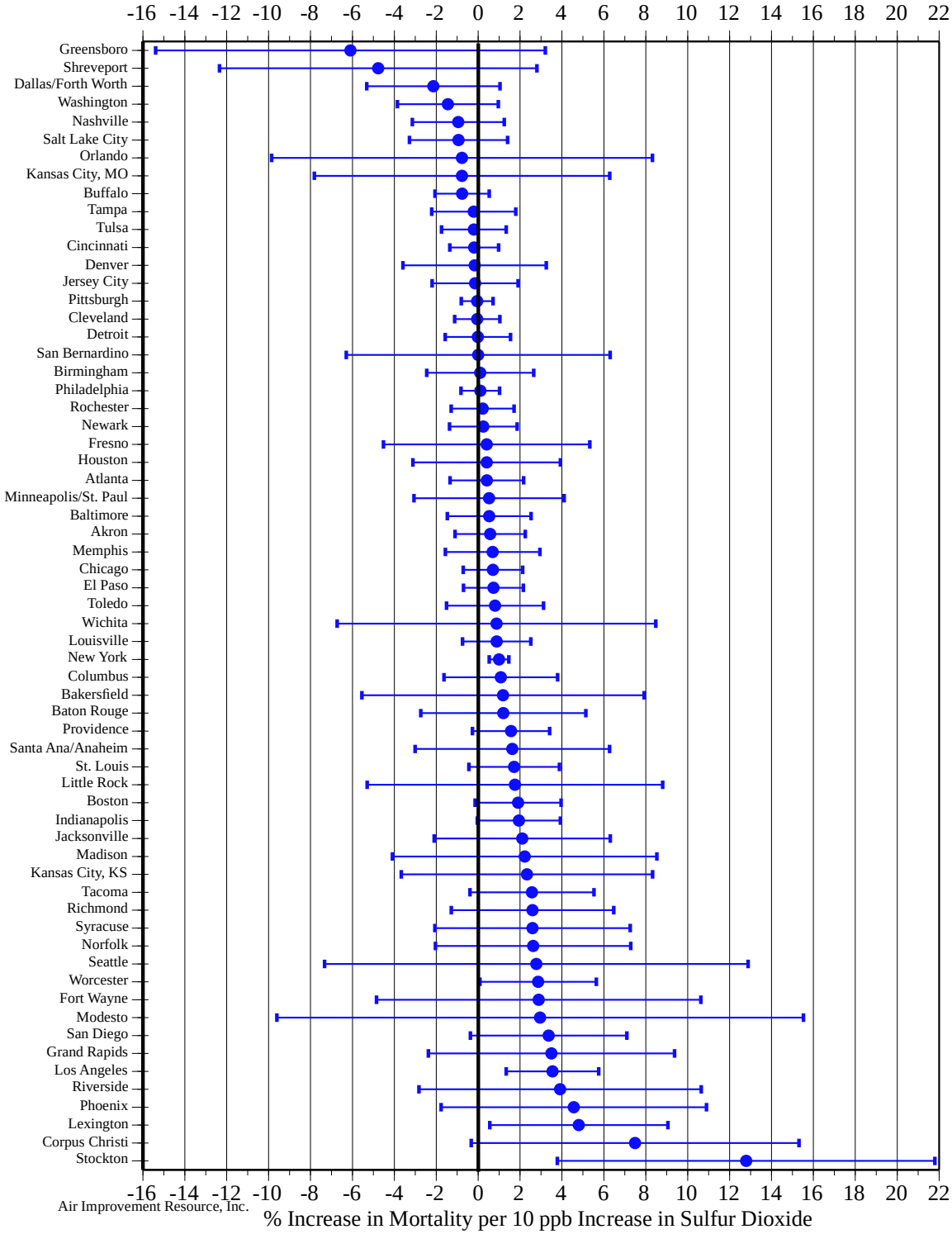


Figure A2
NMMAPS Maximum Likelihood Estimates and 95% Confidence Intervals of the Percentage Increase in Total Mortality from Nonexternal Causes per 10 ppb Increase in Sulfur Dioxide Concentration for Each Location
Lag 1 Day



Air Improvement Resource, Inc.

Figure A3

NMMAPS Maximum Likelihood Estimates and 95% Confidence Intervals of the Percentage Increase in Total Mortality from Nonexternal Causes per 1 ppm Increase in Carbon Monoxide Concentration for Each Location

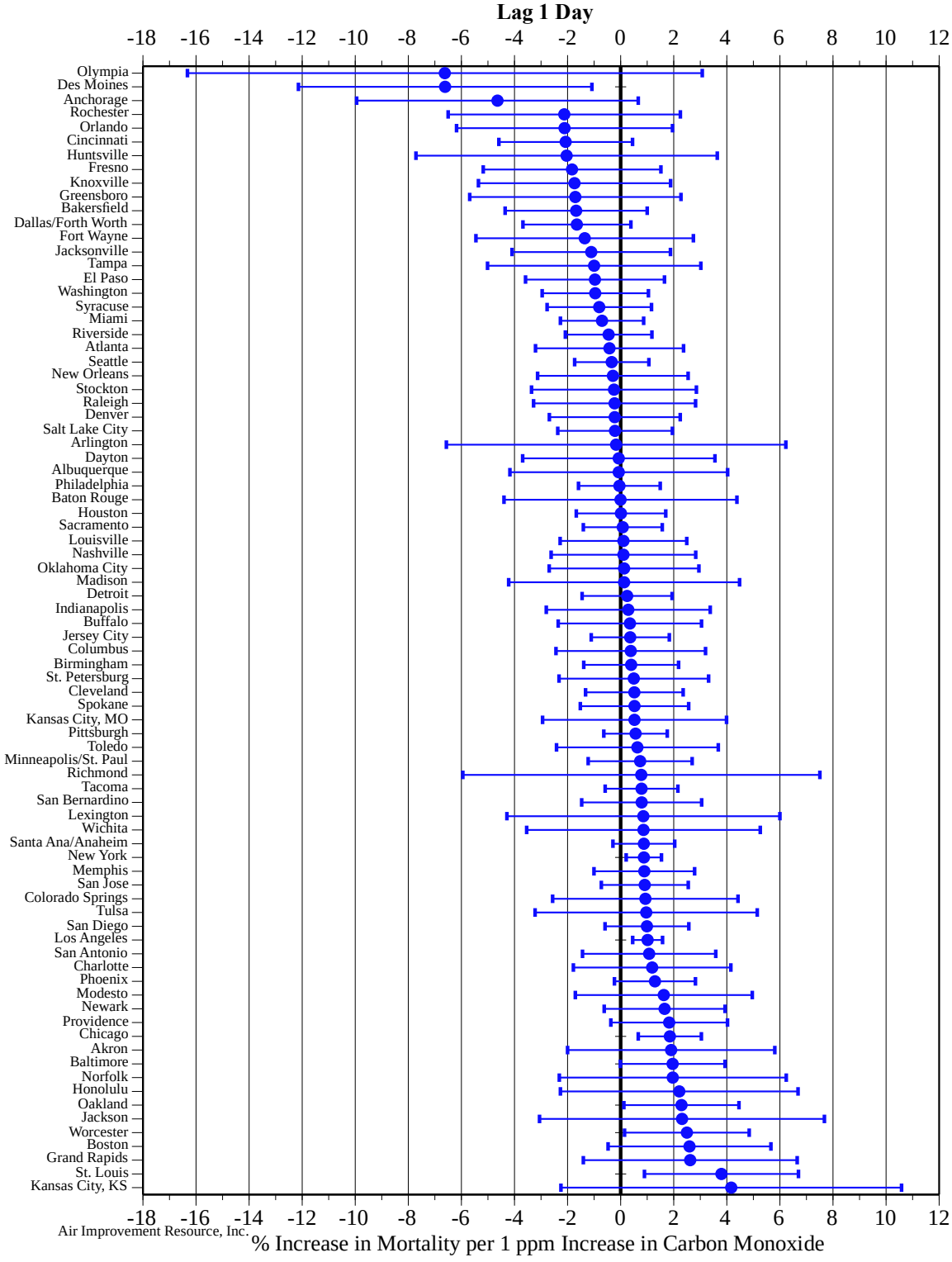


Figure A4
NMMAPS Maximum Likelihood Estimates and 95% Confidence Intervals of the Percentage Increase in Total Mortality from Nonexternal Causes per 10 ppb Increase in Nitrogen Dioxide Concentration for Each Location
Lag 1 Day

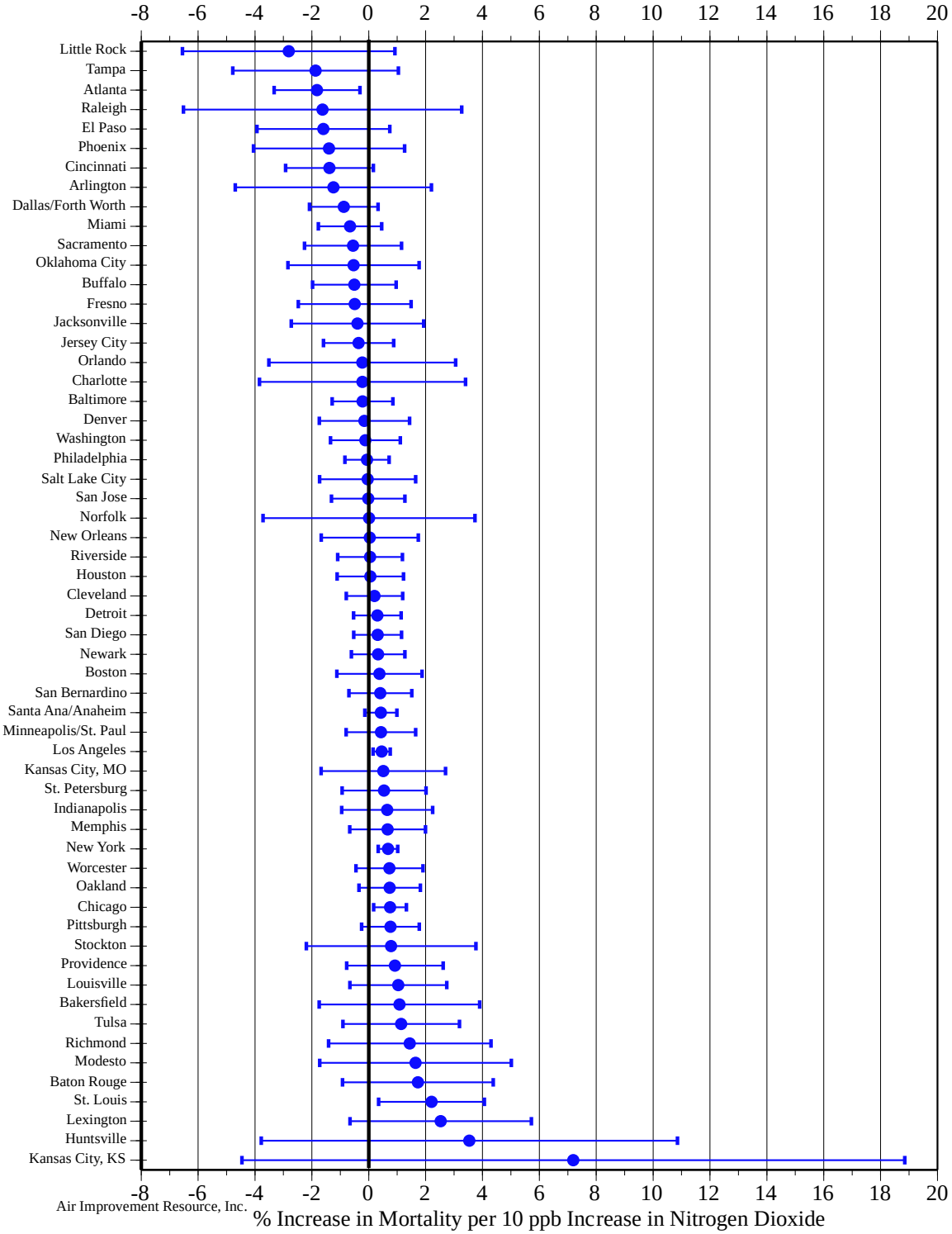


Figure A5

NMMAPS Maximum Likelihood Estimates and 95% Confidence Interval of the Percentage Increase in Total Mortality from Nonexternal Causes per 10 µg/m^3 Increase in PM_{10} Concentration for Each Location

Lag 1 Day

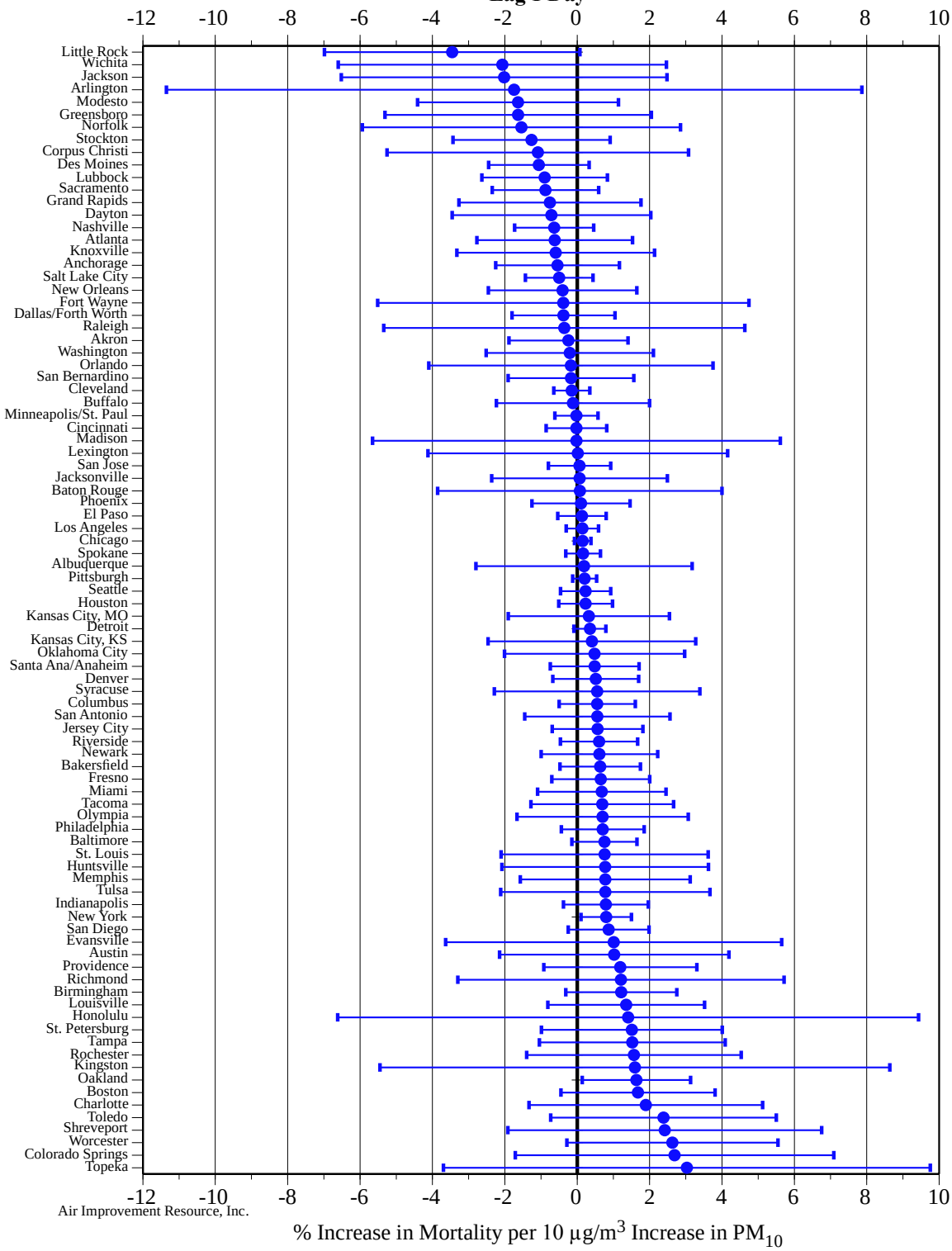


Figure A6
NMMAPS Maximum Likelihood Estimates and 95% Confidence Intervals of the Percentage Increase in Total Mortality from Nonexternal Causes per 10 ppb Increase in Ozone Concentration for Each Location
Lag 1 Day

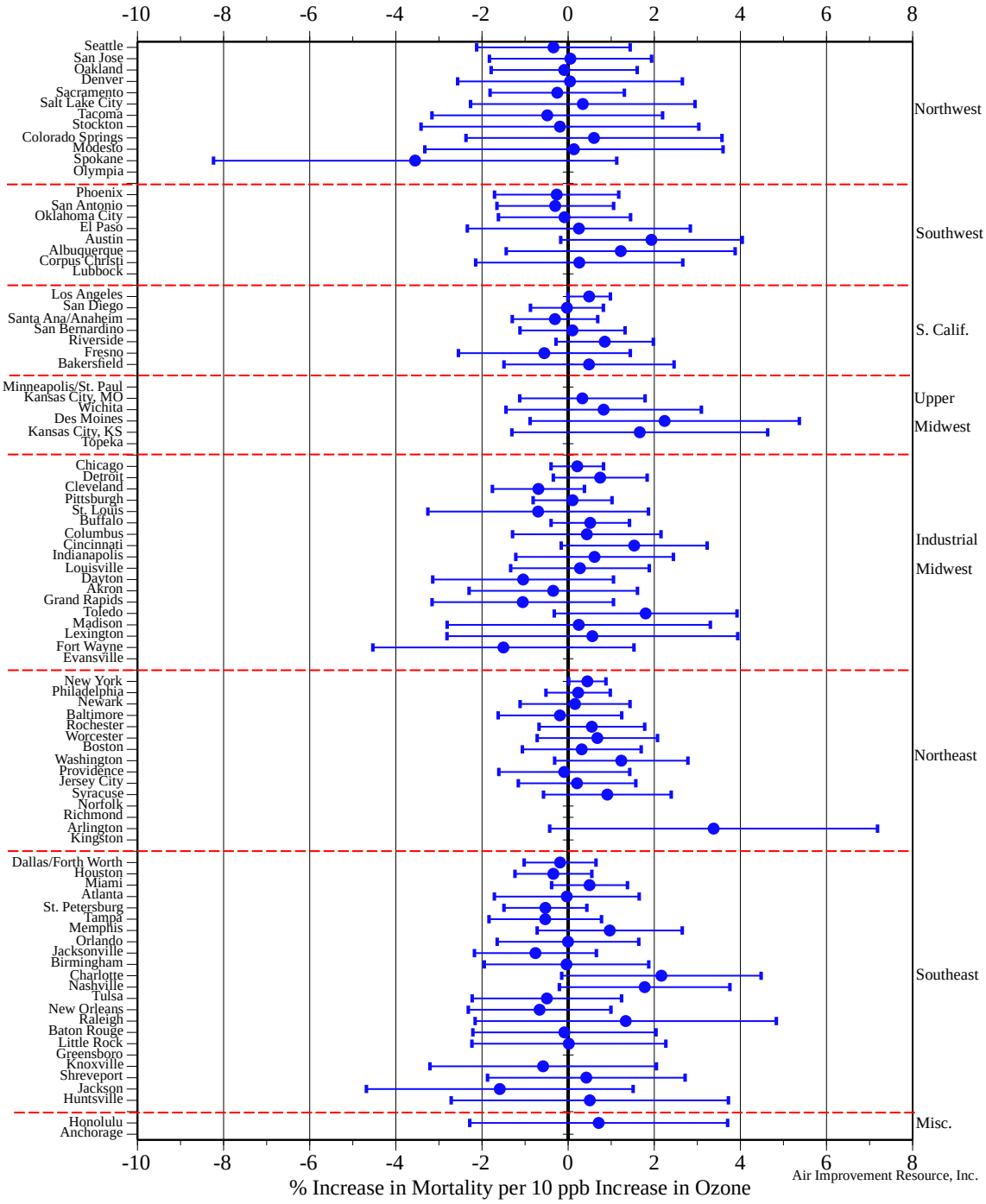


Figure A7
NMMAPS Maximum Likelihood Estimates and 95% Confidence Intervals of the Percentage Increase in Total Mortality from Nonexternal Causes per 10 ppb Increase in Sulfur Dioxide Concentration for Each Location
Lag 1 Day

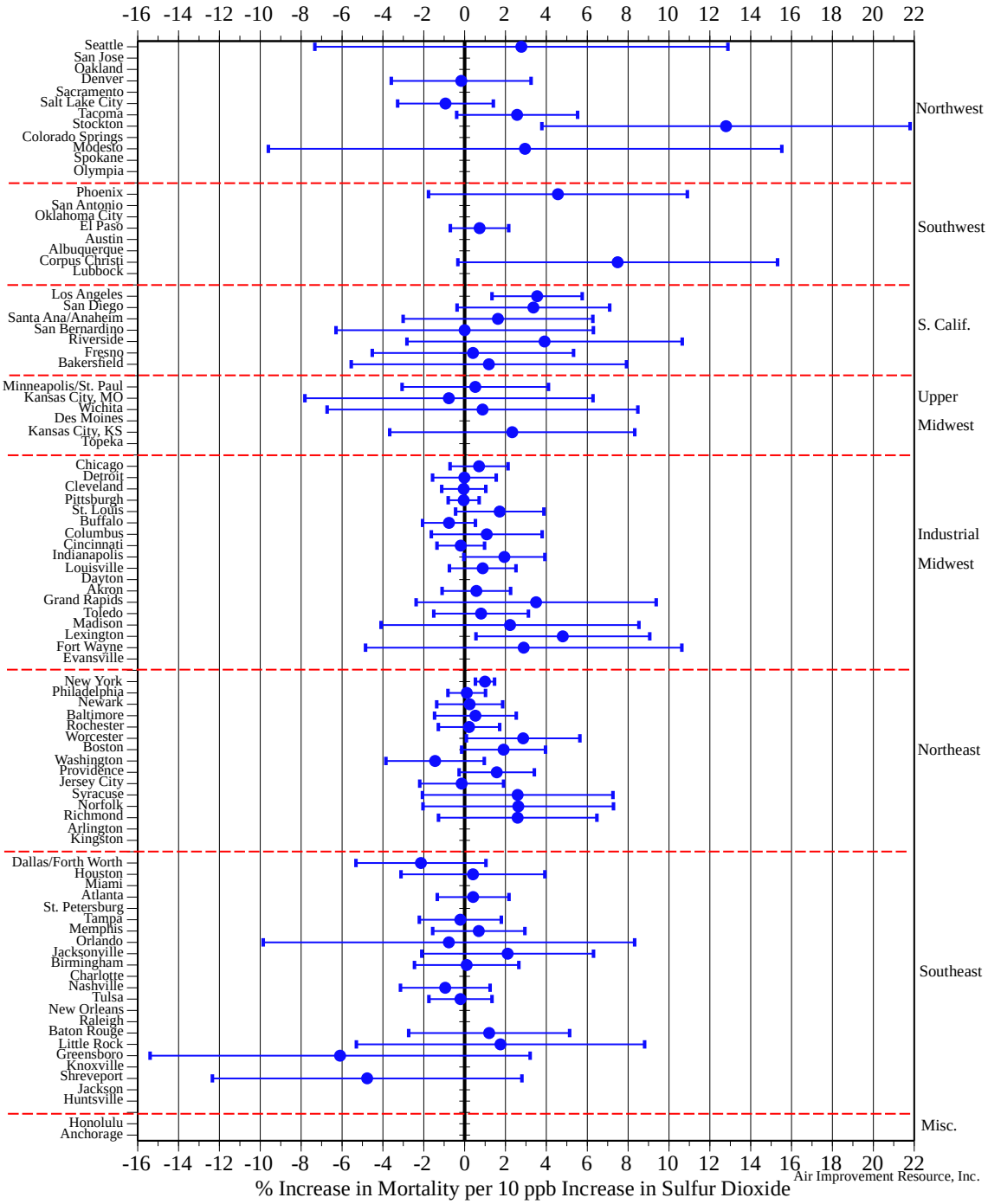


Figure A8
NMMAPS Maximum Likelihood Estimates and 95% Confidence Intervals of the Percentage Increase in Total Mortality from Nonexternal Causes per 1 ppm Increase in Carbon Monoxide Concentration for Each Location
Lag 1 Day

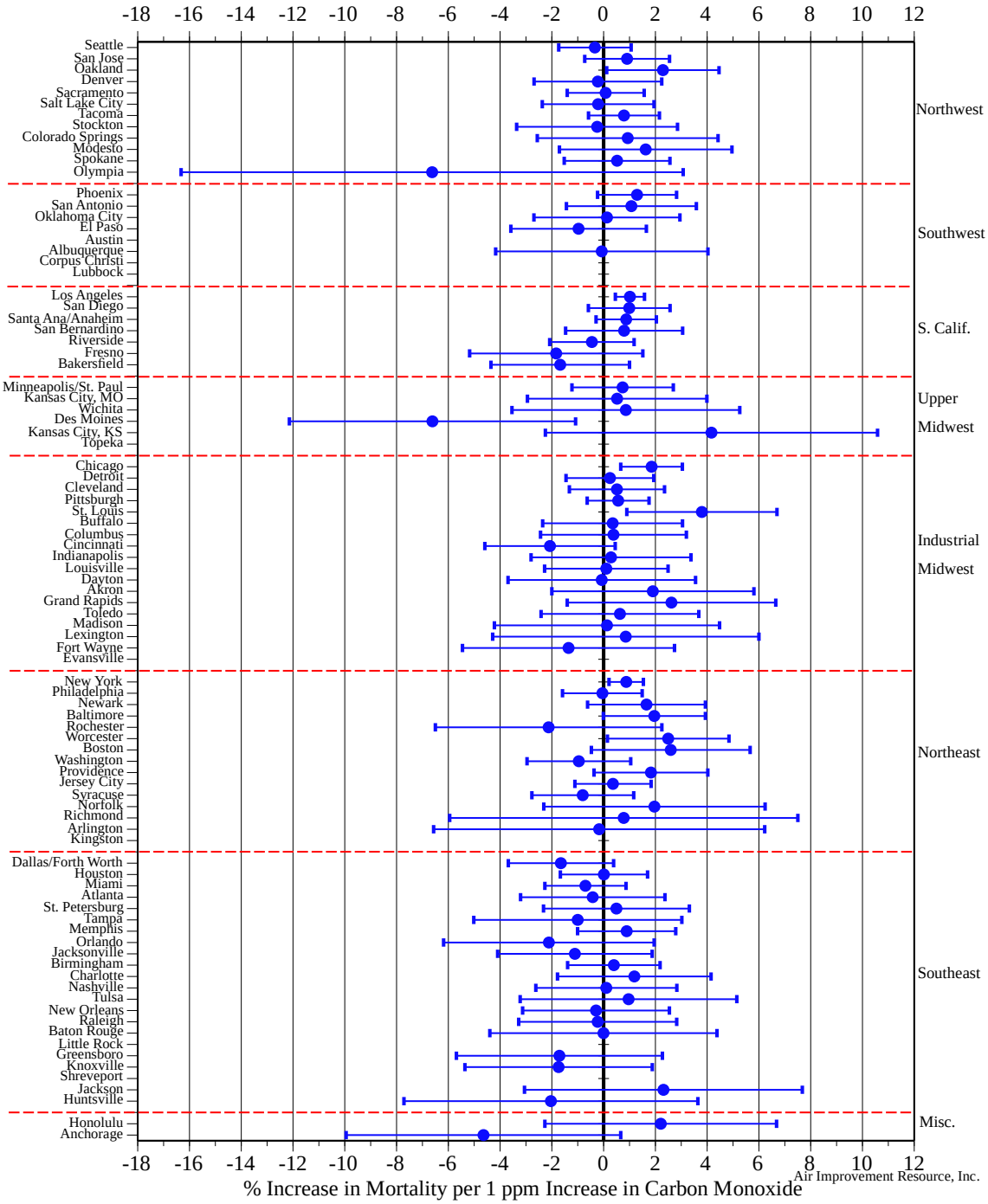


Figure A9
NMMAPS Maximum Likelihood Estimates and 95% Confidence Intervals of the Percentage Increase in Total Mortality from Nonexternal Causes per 10 ppb Increase in Nitrogen Dioxide Concentration for Each Location
Lag 1 Day

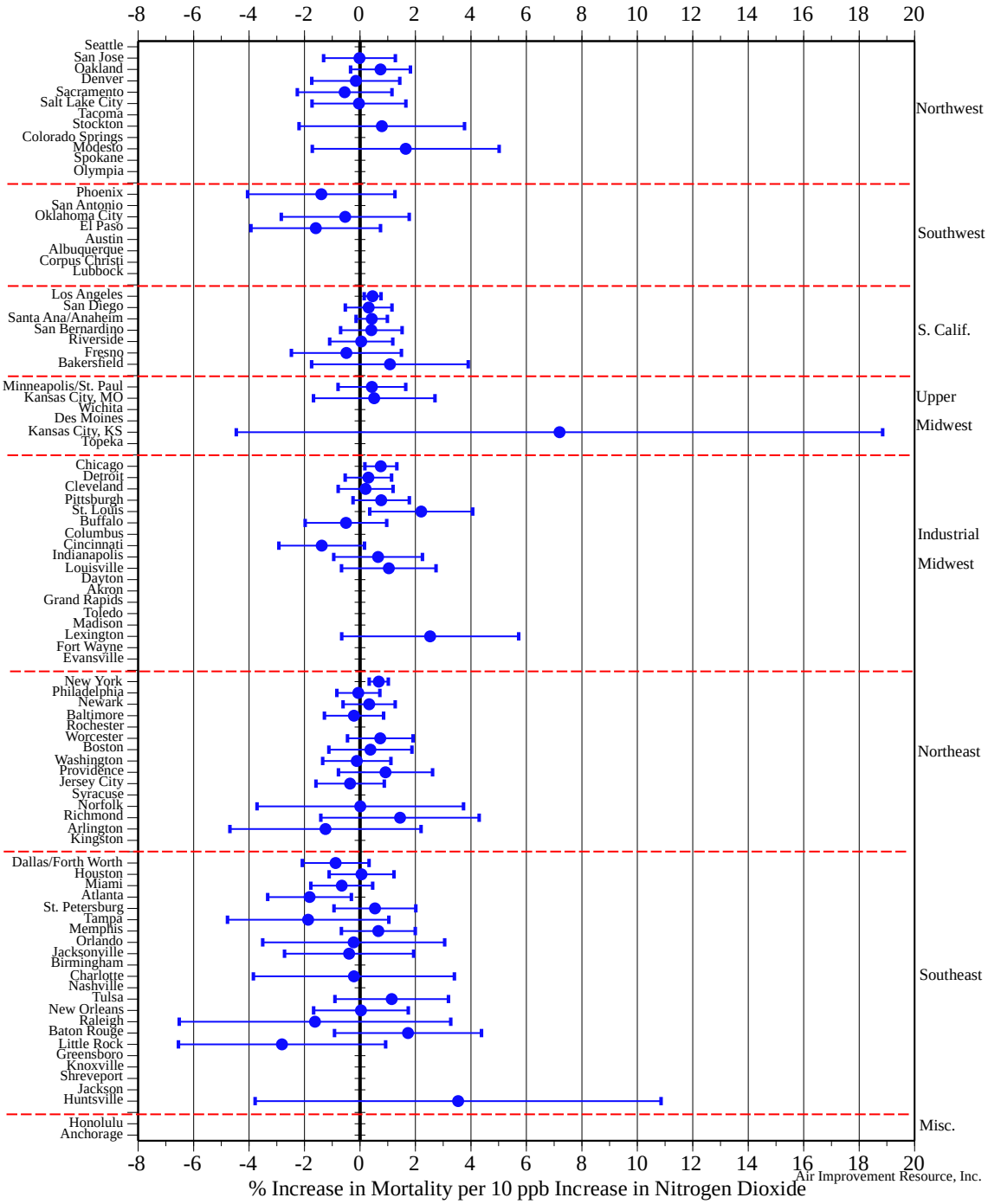
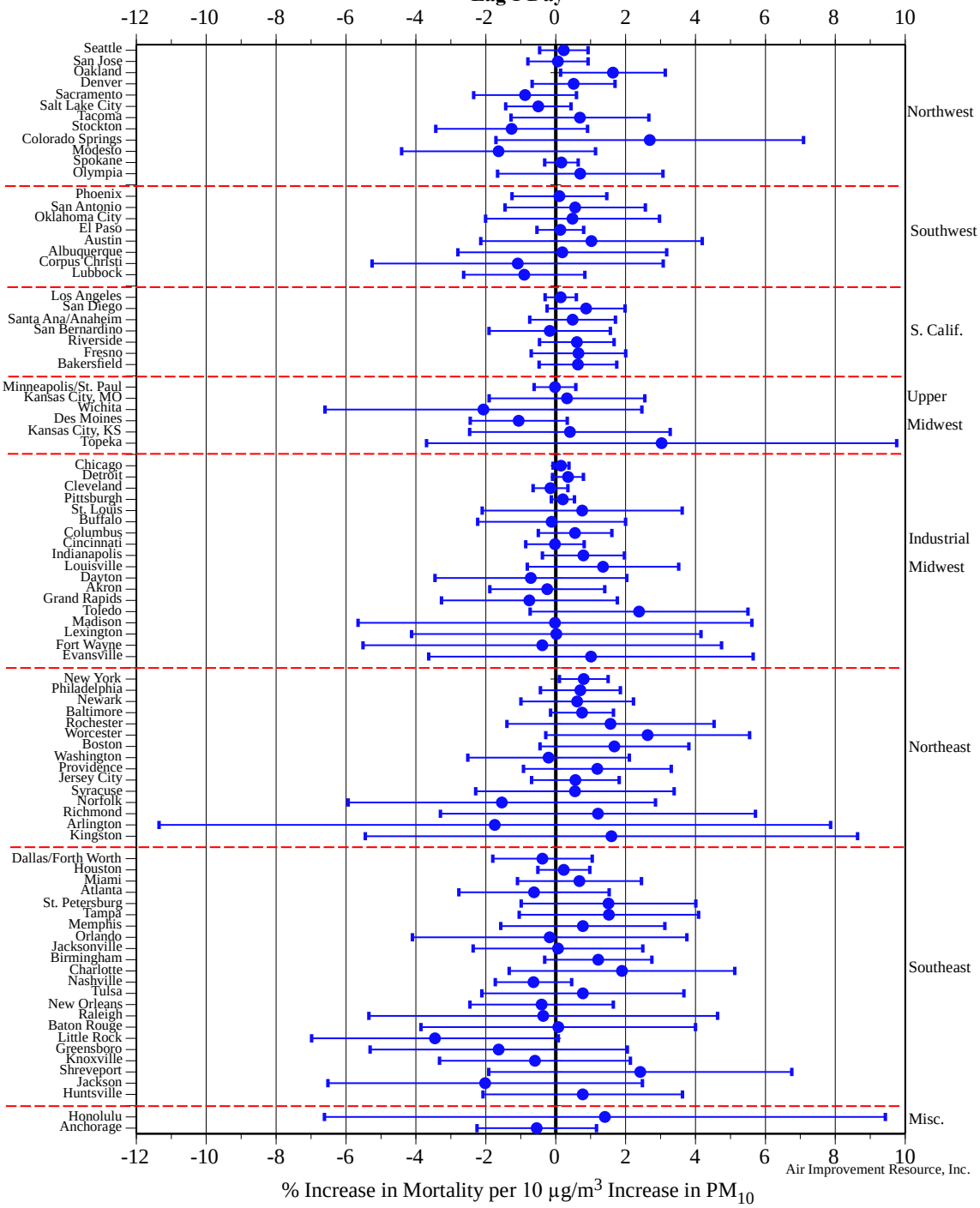


Figure A10
NMMAPS Maximum Likelihood Estimates and 95% Confidence Interval of the Percentage Increase in Total Mortality from Nonexternal Causes per 10 µg/m^3 Increase in PM_{10} Concentration for Each Location
Lag 1 Day



Pollutant	Risk Range	# of Significant + Risks	% +
Ozone	-3.5 to + 3.3	0	53%
Sulfur Dioxide	-6.0 to +13.8	4	71%
Carbon Monoxide	-6.7 to +4.2	5	63%
Nitrogen Dioxide	-2.8 to 7.2	4	60%
PM_{10}	-3.4 to + 3.0	2	63%

Table A1: Summary of the city by city distributions of NMMAPS relative risks for single pollutant models. Column 2 shows the absolute range in risks for the cities. Column 3 shows the number of cities where a statistically significant positive risk was found. Column 4 shows the % of the cities that had a positive risk.