

**Comments on the PM_{2.5} Epidemiology Evaluation in the
Integrated Science Assessment for Particulate Matter
First External Review Draft**

Prepared on behalf of

American Petroleum Institute
1220 L Street, NW
Washington, DC 20005-4070

Prepared by

Julie E. Goodman, Ph.D., DABT
Gradient Corporation
20 University Road
Cambridge, MA 02138

March 12, 2009

Table of Contents

	<u>Page</u>
Executive Summary.....	ES-1
1 Introduction.....	1
2 The framework for causal determination described in Chapter 1 understates uncertainties and is not consistently applied in the ISA.....	2
2.1 Uncertainties are not given consistent or sufficient weight in the evaluation of epidemiology studies in the ISA.....	2
2.1.1 Confounding factors likely accounted for several observed associations.....	3
2.1.2 Measurement error likely biased risk estimates.....	3
2.1.3 Exposure misclassification could have biased results in either direction.....	4
2.1.4 The use of inappropriate statistical models led to biased risk estimates.....	5
2.2 The ISA inappropriately emphasizes non-statistically significant findings in the evaluation of epidemiology studies.....	5
2.3 The Bradford Hill Criteria are used inappropriately in the ISA.....	6
2.3.1 The ISA fails to consider the limitations of studies with weak associations.....	7
2.3.2 The consistency of observed associations is often overstated in the ISA.....	7
2.3.3 Lack of specificity of observed health effects should be given more weight in the ISA.....	7
2.4 The five-level hierarchy for causality presented in the ISA places too much weight on ecological epidemiology studies.....	8
3 Chapter 2 does not accurately portray the epidemiology data as a whole.....	10
3.1 The uncertainty attributed to variations in PM and its components likely has a far greater impact on risk estimates than the ISA suggests.....	10
3.2 Section 2.3 presents a biased portrayal of the weight of evidence from short- and long-term exposure studies of PM _{2.5}	11
3.2.1 Short-term studies do not support a causal association between PM _{2.5} at concentrations below the current NAAQS and health effects.....	12
3.2.2 Long-term studies do not support a causal association between PM _{2.5} at concentrations below the current NAAQS and health effects.....	13
3.3 Section 2.4 does not accurately portray health impacts of PM _{2.5} exposure.....	14
3.3.1 Studies relied on in the ISA to assess concentration-response relationships are not sufficient for concluding a linear model.....	15
4 Studies of short-term PM _{2.5} exposure, such as those reviewed in Chapter 6 of the ISA, do not support a causal association between PM _{2.5} at concentrations below the current NAAQS and health effects.....	17
4.1 Short-term studies of cardiovascular and respiratory morbidity do not provide support for a causal association at levels below the current NAAQS.....	17
4.1.1 Dominici <i>et al.</i> (2006) based risk estimates on exposures measured with central monitors and Medicare data, did not account for co-pollutants, and found results to vary geographically.....	18

4.1.2	Bell <i>et al.</i> , (2008) based risk estimates on exposures measured with central monitors and Medicare data, did not account for co-pollutants, and found results to vary geographically.	19
4.2	Short-term studies of cardiovascular and respiratory mortality do not provide support for a causal association.	21
4.2.1	Franklin <i>et al.</i> (2007) calculated risk estimates that were heterogeneous (and higher in cities with lower PM _{2.5} levels), did not account for co-pollutants, and were sensitive to model selection.	21
4.2.2	Ostro <i>et al.</i> , (2006) calculated risk estimates that appeared to be heterogeneous and were sensitive to co-pollutants and model selection.	22
4.2.3	Burnett <i>et al.</i> (2004) identified NO ₂ as a confounder of PM _{2.5} risk estimates in time-series studies.	24
4.2.4	Dominici <i>et al.</i> (2007) calculated risk estimates based on uncertain exposure estimates that were unadjusted for co-pollutants and were sensitive to model selection.	24
4.3	Selected studies investigating the effects of short-term exposure to PM _{2.5} do not support a causal association with morbidity or mortality at levels below the current NAAQS.	25
5	Studies of long-term PM _{2.5} exposure, such as those reviewed in Chapter 7 of the ISA, do not support a causal association between PM _{2.5} at concentrations below the current NAAQS and health effects.	27
5.1	Long-term studies of cardiovascular morbidity do not provide support for a causal association.	27
5.1.1	Allen <i>et al.</i> (2009) did not report statistically significant risk estimates for subclinical measures of atherosclerosis.	28
5.1.2	Diez Roux <i>et al.</i> (2008) reported mostly null associations (based on linear models) between PM _{2.5} and subclinical measures of atherosclerosis.	29
5.1.3	Hoffman <i>et al.</i> (2006) did not find a sizable influence of background PM _{2.5} on CHD morbidity.	30
5.1.4	Miller <i>et al.</i> (2007) jointly assessed morbidity and mortality, so effects of PM _{2.5} on morbidity alone could not be determined.	32
5.2	Long-term studies of respiratory morbidity do not support a causal association.	32
5.2.1	Goss <i>et al.</i> (2004) based risk estimates on exposures measured with central monitors and did not account for several confounders.	32
5.3	Long-term studies of mortality do not support a causal association.	34
5.3.1	Beelen <i>et al.</i> (2008) reported no association between PM _{2.5} and mortality.	34
5.3.2	Miller <i>et al.</i> (2007) reported high risks (<i>vs.</i> other cohorts) for mortality based on the Cox model, and did not account for co-pollutants or exposure misclassification.	35
5.3.3	Jerrett <i>et al.</i> (2005) calculated risks that were sensitive to model selection and were based on misclassified exposures and on Cox models that did not account for co-pollutants.	36
5.3.4	Laden <i>et al.</i> (2006) calculated risk estimates based on central monitors and used Cox models that didn't account for co-pollutants or other confounders that changed over time.	37
5.3.5	Eftim <i>et al.</i> (2008) based risk estimates on exposures measured with central monitors and Medicare data and did not account for co-pollutants and other confounders.	40

5.3.6	Zeger <i>et al.</i> (2008) based risk estimates on exposures measured with central monitors and Medicare data, did not account for co-pollutants, and found results to vary geographically.....	41
5.3.7	Pope <i>et al.</i> (2009) found a correlation between PM _{2.5} and life expectancy that could have been explained by other factors, such as NO ₂	43
5.4	Long-term studies of reproductive and developmental effects do not provide support for a causal association.	44
5.4.1	Bell <i>et al.</i> (2007) based risk estimates on exposures measured with central monitors and did not rule out NO ₂ or several other confounders.	45
5.4.2	Parker <i>et al.</i> (2008) did not measure PM _{2.5} or exclude preterm births as a result of medical intervention from their study.....	46
5.5	Selected studies investigating the effects of long-term exposure to PM _{2.5} do not support a causal association with morbidity or mortality at levels below the current NAAQS.....	47
6	Chapter 8 does not accurately portray the health impacts of PM _{2.5} exposure.....	49
6.1	Heterogeneity among cities/regions may influence the concentration-response relationship.....	49
6.2	Heterogeneity among the population will not linearize the concentration-response relationship.....	50
6.3	Measurement error can artificially flatten concentration-response curves.	50
6.4	Studies relied on in the ISA to assess concentration-response relationships are not sufficient for concluding a linear model.	51
7	Conclusions	55
	References	56

Executive Summary

In its last review of PM_{2.5}, US EPA (2004) concluded that exposure to ambient PM caused or was associated with a wide variety of health effects, and that no threshold had been identified below which these health effects occur. Based on a review of several recent epidemiology studies, I conclude that the ISA has not adequately demonstrated that studies published since the 2004 US EPA review of PM_{2.5}: (1) demonstrate that PM_{2.5} causes additional health effects not identified in the last review; (2) provide reduced uncertainties or stronger evidence for the previously identified effects; (3) provide evidence that risk estimates for the previously identified effects have increased since the last review; or (4) provide further information on the possibility that these effects occur at lower levels than previously identified.

In Chapter 1 of the ISA, US EPA describes a two-step approach in its framework for causal determination: (1) a five-level hierarchy that considers the Bradford Hill criteria and (2) an evaluation of concentration-response relationships. US EPA's framework understates uncertainties and is not consistently applied in the ISA. Uncertainties – including confounders, measurement error, exposure misclassification, and model uncertainty – are not given consistent or sufficient weight in the evaluation of epidemiology studies. The ISA also inappropriately emphasizes non-statistically significant findings. The ISA uses the Bradford Hill Criteria inappropriately in that it fails to consider limitations of studies with weak associations, the consistency of observed associations is often overstated, and the lack of specificity of exposures and health effects is often not appreciated. Finally, the five-level hierarchy for causality presented in the ISA places too much weight on ecological epidemiology studies, which suffer from many of these uncertainties.

Chapter 2 of the ISA summarizes all of the scientific evidence available since the 2004 AQCD that informs consideration of the policy-relevant questions that frame the ISA. It briefly reviews studies that are discussed in more detail in Chapters 6 and 7, which focus on studies of short-term and long-term PM_{2.5} exposures, respectively. These three chapters do not accurately portray the epidemiological data as a whole. They present a biased portrayal of the weight of evidence from short- and long-term exposure studies of PM_{2.5}, in that uncertainties in the data are not fully considered, very small and non-statistically significant risks are used to support causal determinations, and studies reporting positive associations are emphasized over those reporting no association. Chapter 2 also does not accurately portray the potential health impacts of PM_{2.5} exposure, in that the studies relied on to assess concentration-response relationships are not sufficient for concluding a linear, no-threshold model (discussed in Chapter 8).

In Chapter 6, the ISA concludes that the association between short-term exposure to PM_{2.5} and cardiovascular morbidity is "causal" and that the associations between short-term exposure to PM_{2.5} and respiratory morbidity and mortality are "likely to be causal." In Chapter 7, the ISA concludes that the associations between long-term exposure to PM_{2.5} and cardiovascular morbidity, respiratory morbidity, and mortality are "likely to be causal" and that the association between long-term exposure to PM_{2.5} and reproductive and developmental outcomes is "suggestive." Limitations of several of the major epidemiology studies of short- and long-term exposure to PM_{2.5} published since the 2004 AQCD that were relied on in the ISA do not support these conclusions. For example, the majority of studies reported either null or weakly positive findings. In other cases, weakly positive findings became non-significant when adjusted for confounders. Several studies did not have information on co-pollutants or other factors that may have been associated with exposure and/or outcome, such that reported associations were likely biased away from the null. Exposure misclassification (which could have biased results in either direction) was perhaps the biggest shortcoming of many of the studies considered by US EPA, as almost all studies used measurements from central monitors as surrogates for personal exposures, and other studies did not actually measure exposures at all. Risk estimates across cities within some studies were heterogeneous, yet they were inappropriately pooled for an overall risk estimate. In the short-term exposure studies, risk estimates were often sensitive to the various lag times investigated, with statistically significant associations reported for particular effects at different lag times across the studies. Many long-term exposure studies used the Cox proportional hazard model, which likely led to biased estimates because model assumptions were not always met. More importantly, most long-term exposure studies had few exposures below 15 µg/m³, so they were not informative regarding risks below the current NAAQS.

In Chapter 8, the ISA addresses the concentration-response relationship for PM and key health effects. The ISA presents studies and analyses that suggest the concentration-response relationship for PM and key health effects, in particular morbidity and mortality, may be linear, without a threshold. The ISA does not, however, fully consider evidence that is not consistent with a linear no-threshold concentration-response model for PM and morbidity or mortality.

1 Introduction

Under the Clean Air Act, the United States Environmental Protection Agency (US EPA) is mandated to revise a National Ambient Air Quality Standard (NAAQS) if there is significant new evidence that the standard should be changed. The purpose of an Integrated Science Assessment is to provide the scientific basis for determining whether such evidence exists. The present draft of the ISA for Particulate Matter (PM), First External Review Draft (US EPA, 2008), hereinafter referred to as the "ISA," is not adequate for this purpose.

In the 2004 PM Air Quality Criteria Document (AQCD) (US EPA, 2004), US EPA concluded that exposure to ambient PM caused or was associated with a wide variety of health effects. In addition, US EPA concluded that no threshold had been identified below which these health effects occur. The data presented in the PM ISA do not support that the studies published since the 2004 AQCD: (1) demonstrate that $PM_{2.5}$ causes additional health effects not identified in the last review; (2) provide reduced uncertainties or stronger evidence for the previously identified effects; (3) provide evidence that risk estimates for the previously identified effects have increased since the last review; or (4) provide further information on the possibility that these effects occur at lower levels than previously identified.

Below, I critically review the ISA, in terms of US EPA's criteria for causality and their review of the epidemiological data available since the 2004 AQCD. I also discuss several of the new epidemiology studies of short- and long-term exposure to $PM_{2.5}$ that will likely be key in US EPA's evaluation of whether the standard should be changed. Taken together, these studies do not provide evidence that supports a causal, likely to be causal, or suggestive of a causal relationship for the association between $PM_{2.5}$ and key health effects at exposure levels below the current NAAQS.

2 The framework for causal determination described in Chapter 1 understates uncertainties and is not consistently applied in the ISA.

Section 1.5 of the ISA discusses the US EPA framework for causal determination. The ISA states: "The most compelling evidence of a casual relationship between pollutant exposures and human health effects comes from human clinical studies." The majority of newly available health information evaluated in the ISA comes from epidemiologic studies, although data from human clinical and experimental animal studies are also considered. The ISA states that to move from "association," as reported in epidemiologic studies, to "causation" involves the elimination of alternative explanations for the association. The ISA notes that causal determinations must recognize the uncertainties within scientific data – particularly confounding, measurement error, and exposure misclassification – which are commonly encountered when evaluating health evidence for air pollutants in epidemiologic studies. In the ISA, US EPA uses a two-step approach in their framework for causal determination. The first step uses a five-level hierarchy that classifies the weight of evidence in support of causation, with consideration of the Bradford Hill criteria, and characterizes the strength of any resulting causal classification. The second step evaluates the evidence regarding the concentration-response relationships and the levels, duration, and pattern of exposures at which effects are observed. Although some of the uncertainties within the data are considered in the ISA, the degree to which they might affect the interpretation of risk estimates is often understated. In addition, US EPA does not consistently apply the framework for causal determination across studies or health outcomes in the ISA.

2.1 Uncertainties are not given consistent or sufficient weight in the evaluation of epidemiology studies in the ISA.

Only some of the uncertainties – such as confounding, measurement error, and exposure misclassification – were accounted for in statistical models in *some* of the epidemiologic studies relied upon by the ISA (*e.g.*, Ostro *et al.*, 2006; Goss *et al.*, 2004; Beelen *et al.*, 2008). In addition, some of the statistical models themselves were inappropriate, potentially leading to biased results. Studies that do not account for these uncertainties should be given less weight in the framework for causal determination. Reasons that these uncertainties should be accounted for are described below.

2.1.1 Confounding factors likely accounted for several observed associations.

PM is correlated with many factors, including atmospheric conditions and other co-pollutants. Many of these factors are accounted for in epidemiologic studies of PM, but this is not always the case. Some of the risk estimates from the studies relied on in the ISA were modeled without inclusion of co-pollutants. Even when risk estimates were adjusted for co-pollutants, however, it is still possible that these co-pollutants were not fully accounted for, leading to what is known as "residual confounding," meaning confounding was still present after adjustment (Glymour and Greenland, 2008). If some of the statistically significant risk estimates had fully accounted for confounders, it is possible, and perhaps likely, that they would no longer have been statistically significant.

There may also be unmeasured, or possibly unknown, confounders that could account for observed associations between PM exposure and adverse health effects. These include temperature, humidity, several hazardous air pollutants (HAPs), and stress, for example (Valberg, 2003; Bukowski, 2007, 2008a and b, Goldberg *et al.*, 2008). According to Boffetta *et al.* (2008):

Although the importance of residual confounding and unmeasured confounders as a source of bias in epidemiological studies has been downplayed by many, a recent statistical simulation study showed that with plausible assumptions, effect sizes on the order of 1.5-2.0, which is a magnitude frequently reported in epidemiology studies, can be generated by residual and/or unmeasured confounding.

Although the US EPA considers various confounders in their evaluation of most studies in the ISA, as discussed in greater detail throughout these comments, the bearing of confounders on the interpretation of the results is often not fully appreciated.

2.1.2 Measurement error likely biased risk estimates.

The ISA states that using measurements from central monitors is adequate for assessing human health risks. Yet, estimating individual exposures to air pollutants from central-site outdoor pollution monitors may result in considerable error (Brauer *et al.*, 2002), as discussed in more detail below. Some individuals in the population will have greater exposures than others for any given central-site ambient concentration. This is because exposure measurement error may artificially flatten apparent concentration-response curves and tend to make any concentration-related effect (even those that are truly threshold in nature) look more or less linear as an artifact of the analysis, thus masking what may in fact

be a steeper curve (Brauer *et al.*, 2002; Rhomberg, 2009). The possibility that exposure measurement error obscures thresholds limits the ability to draw conclusions about effects of PM_{2.5} at low exposure levels.

2.1.3 Exposure misclassification could have biased results in either direction.

In studies in which exposure is measured in categories (*e.g.*, quartiles, quintiles), exposure misclassification can result when concentrations measured at central monitors are not representative of personal exposures. Reasons for this include uneven distribution of PM attributable to local sources; monitoring sites may represent a nearby source and not human exposures a small distance away; pollution patterns can be affected by terrain features and weather; and daily variations in PM concentrations at a central monitoring site may differ from variations experienced by individuals. These factors may bias the results of an epidemiological analysis in either direction.

Exposure misclassification for PM is likely to be non-differential. Non-differential misclassification means that every subject, regardless of disease status, has an equal chance of being misclassified because which subjects are misclassified is a matter of chance. The actual *fraction* of subjects in a particular study misclassified in the diseased and non-diseased groups is likely to be different. Even if misclassification is non-differential on average, due to random variation, misclassification rates in a single study will most likely be differential (Jurek *et al.*, 2005; 2008), and may bias results in any direction. In fact, Sorahan and Gilthorpe (1994) showed that a considerable percentage of studies with non-differential misclassification produced risk estimates that were larger than those from data sets that were classified correctly. According to Wacholder *et al.* (1995):

Several papers published since 1990 have shown that there are special circumstances where there is a bias towards exaggeration of effects. Dosemeci *et al.* identified a scenario where non-differential misclassification of exposure more often than not leads to an overestimate of the odds ratio in an intermediate exposure category when there are more than two exposure levels. Other papers that have appeared since the textbooks cited by Sorahan and Gilthorpe' were published during the 1980s, have identified circumstances where an overestimate is more likely than an underestimate. These include particular forms of non-differential misclassification when an exposure is not binary, when grouping has occurred, or when the errors in a continuous exposure are correlated with their true value.

Because of the high likelihood of exposure misclassification and the impossibility of knowing with certainty in which way this will bias results, the epidemiology data relied on in the ISA are insufficient to determine whether PM exposure is associated with health effects at low-level exposures.

2.1.4 The use of inappropriate statistical models led to biased risk estimates.

Every risk estimate is highly dependent on the statistical model from which it is calculated. If a model is based on assumptions that are not met, then the risk estimate is likely to be biased. For example, Moolgavkar (2005) suggested that the assumptions of the Cox proportional hazards model are violated in many ecological studies of pollution and health effects. This is likely the case for several long-term PM_{2.5} exposure studies referenced in the ISA (*e.g.*, Beelen *et al.*, 2008; Miller *et al.*, 2007; Jerrett *et al.*, 2005; Laden *et al.*, 2006). As stated by Abrahamowicz *et al.* (2003):

[T]he proportional hazards (PH) assumption... implies that the impact of each covariate on hazard remains constant during the entire follow-up time. While testing the PH assumption is interesting in its own right, simultaneous modeling of nonlinear and time-dependent effects of the exposure of interest may be necessary to avoid biased estimates and incorrect conclusions. (Abrahamowicz *et al.*, 1996)

This means that not only the impacts of exposure, but also those of all potential confounders, must be proportional over time to prevent a biased risk estimate. Abrahamowicz *et al.* (2003) actually tested whether this held for a subset of the American Cancer Society (ACS) Cancer Prevention Study II, which had PM_{2.5} data for 50 cities and sulfate data for 151 cities. They examined the effects of PM_{2.5} and sulfate on all-cause mortality in a sub-cohort of 1,200 individuals and 1,300 cases (*i.e.*, deaths) by pooling the results of separate analyses of 10 disjoint random subsets of the entire dataset, each with ~2,200 participants. They found for both PM_{2.5} and sulfate, there was a statistically significant deviation from the traditional linearity assumption. This was also true of body mass index (BMI), a confounder in the model. Based on a flexible regression spline generalization of the Cox proportional hazard model, which was not restricted to the same assumptions of the usual Cox PH model, they found that risk estimates for both PM_{2.5} and sulfate differed from those based on models using the traditional assumptions. While risks for PM_{2.5} were inflated at low doses, sulfate was shown to have a threshold. This demonstrates that Cox PH models do not give accurate risk estimates, particularly at low doses.

2.2 The ISA inappropriately emphasizes non-statistically significant findings in the evaluation of epidemiology studies.

The ISA states in Section 1.5.2: "Much of the newly available health information evaluated in this ISA comes from epidemiologic studies that report a statistical association between ambient exposure

and health outcome." Tests of statistical significance are an important way to reach a conclusion about a sample population, while quantifying the chance that the conclusion is incorrect. Commonly, the word significant is used to indicate importance; however in statistical analysis, it refers to the probability that an observed effect or relationship is true (*i.e.*, not a result of random variability), given certain assumptions. Thus, one should have a lower degree of confidence that there are true differences between groups if the differences in measured outcomes are not statistically significant.

In the ISA, non-statistically significant effects were often considered as evidence of an association. If, within and among studies, associations are consistently in one direction but not statistically significant (perhaps because individual studies do not have enough power for statistically significant results to be observed), it is possible in certain circumstances that together these associations are indicative of an effect. This is not always the case, however, and in general it is scientifically inappropriate to interpret results that are not statistically significant as suggestive evidence of a health effect. This is particularly true when there are non-significant associations in two directions; the associations in the positive direction should not be given more weight than those in the negative direction. In addition, many small, non-significant associations may be more indicative of a consistent confounder than a true association. In general, only associations that are statistically significant should be considered "positive" in a weight-of-evidence analysis for causation, and non-significant findings should be interpreted with caution.

2.3 The Bradford Hill Criteria are used inappropriately in the ISA.

In the ISA, US EPA modified the Bradford Hill criteria (Hill, 1965) for use in causal determinations specific to health effects of pollutant exposures, such that they can be used with data from epidemiologic, controlled human exposure, and animal toxicological studies. The ISA states that the criteria that usually play a larger role in the determination of causality are consistency of results across studies, coherence of effects observed in different study types or disciplines, biological plausibility, the exposure-response relationship, and evidence from "natural" experiments (epidemiological studies). Other criteria that US EPA considers regarding epidemiological data are the strength, specificity, and temporal relationship of the observed associations. US EPA routinely uses the Bradford Hill criteria inappropriately in the ISA, however, as described below.

2.3.1 The ISA fails to consider the limitations of studies with weak associations.

Regarding the criterion of strength of observed associations, the ISA does not give appropriate weight to the magnitude of the risk estimates, almost all of which are very small. In the ISA, all risk estimates are considered positive and supportive of a causal association if they are above null, regardless of whether they are statistically significant or not. Yet, many of the small, statistically significant associations are found to be non-significant when confounders are accounted for. As discussed above, it is likely that residual, unmeasured, and/or unknown confounders could have accounted for many of the observed associations.

2.3.2 The consistency of observed associations is often overstated in the ISA.

The reproducibility of findings within and across studies is one of the strongest arguments for a true association and, thus, causality. In the ISA, the "consistency" of the results is routinely overstated. Part of this is because almost all risk estimates greater than 1 are considered "positive" in the ISA, regardless of statistical significance. Negative or null associations are rarely mentioned and do not seem to be considered in the overall weight of evidence.

Another issue is that if there are confounders consistently found in all of the studies, this will also lead to consistent findings, but they will all be at least partially attributable to confounders and consistently biased away from the null. As discussed above, there are many known confounders, and likely some unknown, that may have led to observed associations in the PM studies. As stated by Boffetta *et al.* (2008):

There is a fundamental difference between false positive results that are generated by chance and those caused by bias – the former will rarely be replicated in subsequent investigations, whereas bias may operate in a similar fashion in different settings and populations and thus will provide a consistent pattern of independently generated results. Even if only a relatively low proportion of results are generated by bias, the probability of false-positive discovery may be substantial.

2.3.3 Lack of specificity of observed health effects should be given more weight in the ISA.

Specificity can be examined by determining whether one disease is specific to one agent or whether one agent is specific to one disease. In the ISA, specificity is considered to be one of the weaker guidelines for causality, because many agents or other factors (e.g., other pollutants, stress) can cause

certain health effects associated with PM. This is an important issue for respiratory and cardiovascular morbidity and mortality, however, because some of these other causal factors are correlated with PM exposure. For example, Bukowski (2007; 2008a,b) and (Goldberg *et al.* 2008) discuss correlations among driving stress and roadway pollutant exposure. The possibility that some of effects attributed to PM may be linked to these other factors has not been adequately considered by US EPA in the ISA. Moreover, the ISA does not discuss the second aspect of specificity; that is, whether one agent is specific to one disease. The ISA often combines several different types of health effects (*e.g.*, cardiovascular and respiratory morbidity, hospital admissions and other measures of morbidity). While it is certainly possible that one agent can cause several different health effects, one must determine whether there is a biological basis for combining them when conducting causation analyses. In addition, it should be stressed that an increased risk of a particular health effect in one study and another type of health effect in another study is not consistent evidence of an effect. The ISA should consider the lack of specificity with respect to both other contributing factors for health effects associated with PM and the different types of health effects associated with PM in its causal determinations.

2.4 The five-level hierarchy for causality presented in the ISA places too much weight on ecological epidemiology studies.

In the ISA, US EPA uses a five-level hierarchy that classifies the weight of evidence for causation, not just association; that is, whether the weight of scientific evidence makes causation at least as likely as not. The five classifications are: "Causal relationship," "Likely to be a causal relationship," "Suggestive of a causal relationship," "Inadequate to infer a causal relationship," and "Suggestive of no causal relationship."

The ISA states that the five-level hierarchy is based on the scheme used in the Institute of Medicine's *Improving the Presumptive Disability Decision-Making Process for Veterans* (IOM, 2008). The IOM used a four-level hierarchy, with classifications of the scientific evidence as: "Sufficient" to conclude that a causal relationship exists; "Equipose and Above," in which the evidence is sufficient to conclude that a causal relationship is at least as likely as not; "Below Equipose," in which the evidence is not sufficient to conclude that a causal relationship is at least as likely as not, or is not sufficient to make a scientifically informed judgment; and "Against" a causal relationship. The "Equipose and Above" classification is used when the preponderance of epidemiologic studies show evidence of an association that cannot readily be explained by uncertainties such as chance, bias, or confounding, whereas the

"Below Equipose" classification is used when the epidemiologic evidence is limited by the inability to rule out these uncertainties. Ecological epidemiology studies are much more limited by uncertainties (*e.g.* exposure misclassification) compared to studies that use individual exposure-level data. The five-level hierarchy used in the ISA does not draw a distinction between epidemiology studies based on group-level data and those based on individual-level data. The "suggestive" criterion places too much confidence in the results of ecological epidemiology studies, and the ISA states that this criterion is met if "at least one high quality study shows a positive association but the results of other studies are inconsistent." The results of ecological studies are often inconsistent and, therefore, one positive study combined with several inconsistent studies should be considered inadequate to infer a causal relationship. This is particularly true because even in most "high quality" ecological studies, bias, chance, or confounding by other pollutants cannot be completely ruled out.

3 Chapter 2 does not accurately portray the epidemiology data as a whole.

Chapter 2 of the ISA summarizes the newly available scientific evidence (since the 2004 AQCD) that informs consideration of the policy-relevant questions that frame the ISA. This summary, however, does not accurately portray the data as a whole. It presents a biased portrayal of the weight of evidence from short- and long-term exposure studies of PM_{2.5}, in that uncertainties in the data are not fully considered, very small and non-statistically significant risks are used to support causal determinations, and studies reporting positive associations are emphasized over those reporting no association. It also does not accurately portray the health impacts of PM_{2.5} exposure at levels below the current NAAQS, in that the studies relied on to assess concentration-response relationships are not sufficient for concluding a linear, no-threshold model.

3.1 The uncertainty attributed to variations in PM and its components likely has a far greater impact on risk estimates than the ISA suggests.

Section 2.1 discusses the trends in ambient concentrations and sources of PM. In this section, the ISA acknowledges the inadequacies of US EPA Air Quality System (AQS) data, which can be a source of uncertainty that can affect risk estimates in studies relied on in the ISA. These inadequacies include that a majority of US counties are not represented because of population density requirements, monitors are not uniformly distributed across counties or across the US, and spatial variability is likely region-specific and influenced by local sources and conditions.

Section 2.2 discusses the evidence regarding personal exposure to ambient PM in outdoor and indoor microenvironments and the relationship between ambient PM concentrations and exposure to PM from ambient sources. The ISA notes that for outdoor exposure, monitored PM concentrations and true community average concentrations are affected by many factors, such as monitor location and height, wind direction, and terrain. In addition, the spatial and temporal mobility of humans results in personal concentrations that differ from those obtained at a central site. For indoor exposure, infiltration factors vary depending on particle size, region, and season; concentration also depends on building ventilation properties and practices. PM components can also be heterogeneous across regions. For example, sulfate exposure is higher in the eastern than the western US, and vehicle emissions and secondary nitrate exposures are higher in the western than in the eastern US.

Thus, variations of PM and its components can be a source of uncertainty when using PM monitoring at a central site as a surrogate for exposure. US EPA recognizes this in the ISA, but states that risk estimates based on ambient concentrations are an appropriate measure for risk management purposes, even though they increase the standard error, because they give the change in health effects "resulting from" a change in ambient concentration of PM. Such a statement assumes, incorrectly, that a correlation between concentrations and the change in health effects indicates causation. Also, the sources of PM and the chemistry of the PM mixture that are most important for determining health effects are unclear. As discussed below, the uncertainty attributed to variations in PM and its components likely has a far greater impact on risk estimates than the ISA suggests.

3.2 Section 2.3 presents a biased portrayal of the weight of evidence from short- and long-term exposure studies of PM_{2.5}.

Section 2.3 summarizes and classifies the health effects of short- and long-term exposure to PM_{2.5}. These classifications are based on epidemiological, clinical, and toxicological studies, when available. Although the criteria for causality are set out in Chapter 1 of the ISA, they are somewhat subjective and not always uniformly applied to these outcomes.

In the ISA, it was determined that a causal relationship exists between short-term exposure to ambient concentrations of PM_{2.5} and cardiovascular morbidity, that a causal relationship is likely to exist between short-term exposure to ambient concentrations of PM_{2.5} and respiratory morbidity and mortality, and that a causal relationship is likely to exist between long-term exposure to ambient concentrations of PM_{2.5} and cardiovascular morbidity, respiratory morbidity, and mortality.

In order to provide a useful representation of the recent science, an unbiased assessment of that science must be made. US EPA has not made such an assessment in the ISA. For example, the ISA does not consistently give sufficient weight to uncertainties, exposure misclassification, and confounders, all of which may bias results. In addition, the ISA focuses on positive associations whether or not they are statistically significant. In the ISA, it is assumed that studies reporting very small and non-statistically significant risks can be used to support causal determinations, as well as the hypothesis that there is no threshold below which health effects occur. In many "positive" studies, there was a focus on the single or few results that were statistically significant among the many available results, such as one lag time or age group among many, or one model among many, or results for one city with positive results when other

cities were negative. It is scientifically inappropriate to interpret results that are barely above null and not statistically significant as suggestive evidence of a health effect.

3.2.1 Short-term studies do not support a causal association between PM_{2.5} at concentrations below the current NAAQS and health effects.

The ISA concludes that the association between short-term exposure to PM_{2.5} and cardiovascular morbidity is "causal" and that the associations between short-term exposure to PM_{2.5} and respiratory morbidity and mortality are "likely to be causal." I reviewed two studies upon which US EPA based its conclusions regarding short-term PM_{2.5} exposure and cardiovascular and respiratory morbidity (Dominici *et al.*, 2006; Bell *et al.*, 2008) and four studies that examined short-term PM_{2.5} exposure and mortality (Franklin *et al.*, 2007; Ostro *et al.*, 2006; Burnett *et al.*, 2004; Dominici *et al.*, 2007). Based on the review of these short-term exposure studies, I conclude that the new epidemiology studies do not support a causal association between PM_{2.5} at concentrations below the current NAAQS and health effects. That is, these studies should not be used as evidence of a "causal" or "likely causal" relationship between short-term exposure to ambient PM_{2.5} and cardiovascular morbidity, respiratory morbidity, or mortality.

Several of these studies reported either null or weakly positive findings (Dominici *et al.*, 2006; Bell *et al.*, 2008, Franklin *et al.*, 2007). In other cases, weakly positive findings became non-significant when adjusted for confounders (Ostro *et al.*, 2006; Burnett *et al.*, 2004). Several studies did not have information on co-pollutants, so reported associations were likely biased away from the null (Dominici *et al.*, 2006, 2007; Bell *et al.*, 2008, Franklin *et al.*, 2007). Risk estimates across cities were heterogeneous in many of these studies (Dominici *et al.*, 2006; Bell *et al.*, 2008, Franklin *et al.*, 2007); so pooling them was inappropriate. Risk estimates in most of the studies were sensitive to the various lag times investigated (Dominici *et al.*, 2006, 2007; Bell *et al.*, 2008, Franklin *et al.*, 2007; Ostro *et al.*, 2006). Moreover, statistically significant associations were reported for particular effects at different lag times across the studies. For example, Dominici *et al.* (2006) reported a statistically significant increased risk for COPD at lag 0, but Bell *et al.* (2008) reported a statistically significant increased risk for this same outcome only at lag 2.

Exposure misclassification/measurement error is perhaps the biggest shortcoming of the short-term studies of PM_{2.5}. All of the studies reviewed here used measurements from central monitors as surrogates for personal exposure. The distance between people's residences to these monitors can vary, increasing the likelihood of obtaining inaccurate measurements.

All of these factors make it difficult to attribute risks to short-term exposure to PM_{2.5}, and several of these studies did not report actual exposures (Bell *et al.*, 2008; Dominici *et al.*, 2007). Thus, these short-term studies are not informative regarding risks below the current NAAQS.

3.2.2 Long-term studies do not support a causal association between PM_{2.5} at concentrations below the current NAAQS and health effects.

In the ISA, the conclusion by US EPA that there is a likely causal relationship between long-term exposure to PM_{2.5} and cardiovascular morbidity appears to be based on research of short-term exposure. The ISA states that the results were inconsistent in long-term exposure studies, but "the evidence from epidemiologic, human clinical, and animal toxicological studies that examined the cardiovascular outcomes associated with short-term exposure to PM_{2.5} (discussed in Section 6.2), supports a role for the development of cardiovascular morbidity in response to long-term exposure to PM_{2.5}. Based on the consistent and coherent evidence from epidemiologic and toxicological studies that examined the association between long-term and short-term exposure to PM_{2.5} and cardiovascular morbidity, sufficient evidence is available to conclude that a causal relationship is likely to exist between long-term exposure to ambient concentrations of PM_{2.5} and cardiovascular morbidity." The ISA states that the association between PM_{2.5} and respiratory morbidity is likely to be causal because "collectively," toxicological studies provide biological plausibility and "overall," evidence from epidemiological and toxicological studies is consistent and coherent. The ISA also states: "The new epidemiologic evidence reports a consistent association between long-term exposure to PM_{2.5} and an increased risk of mortality (with the majority of the effects ranging from > 1 to 1.20) in cities with annual average PM_{2.5} concentrations ranging from 10.2-29 µg/m³ (see Section 7.6)."

I reviewed four studies on which US EPA based its conclusions regarding long-term PM_{2.5} exposure and cardiovascular morbidity: two of which analyzed subclinical effects (Allen *et al.*, 2009; Diez Roux *et al.*, 2008), and two of which analyzed clinical outcomes in epidemiological studies (Hoffman *et al.*, 2006; Miller *et al.*, 2007). I also reviewed one study, by Goss *et al.* (2004), that US EPA relied on for its assessment of respiratory morbidity, seven of the major studies published since the 2004 AQCD that examined PM_{2.5} exposure and mortality (Beleen *et al.*, 2008; Eftim *et al.*, 2008; Jerrett *et al.*, 2005; Laden *et al.*, 2006; Miller *et al.*, 2007; Pope *et al.*, 2009; Zeger *et al.*, 2008), and two studies relied on in the ISA for assessment of reproductive and developmental outcomes (Bell *et al.*, 2007; Parker *et al.*, 2008).

Based on my review of these long-term exposure studies, I conclude that the new epidemiology studies do not support a causal association between PM_{2.5} at concentrations below the current NAAQS and health effects. That is, the long-term exposure studies of PM_{2.5} should not be used as evidence of a "likely" causal relationship between ambient PM_{2.5} at concentrations below the current NAAQS and cardiovascular morbidity, respiratory morbidity, or mortality, nor as "suggestive" evidence of a causal relationship between PM_{2.5} below the NAAQS and reproductive and developmental outcomes.

Several of these studies reported either null or weakly positive findings (Allen *et al.*, 2009; Diez Roux *et al.*, 2008). In other cases, weakly positive findings became non-significant when adjusted for confounders (*e.g.*, Allen *et al.*, 2009). Several studies did not have information on co-pollutants or other factors that may have been associated with exposure and/or outcome (such as people living closer to monitors may have low SES and be at higher risks for certain outcomes), so reported associations were likely biased away from the null (Allen *et al.*, 2009; Diez Roux *et al.*, 2008; Beelen *et al.*, 2008; Eftim *et al.*, 2008; Jerrett *et al.*, 2005; Laden *et al.*, 2006; Miller *et al.*, 2007; Pope *et al.*, 2009; Zeger *et al.*, 2008).

Exposure misclassification/measurement error is perhaps the biggest shortcoming of long-term studies of PM_{2.5}. All studies reviewed here used measurements from central monitors and, because the distance between people's residences to these monitors varied, this lead to inaccurate measurements. In addition, some studies used exposure measurements from 2000 to represent earlier exposures (*e.g.*, Jerrett *et al.*, 2005). As PM_{2.5} concentrations have been decreasing over time, this likely overestimated risks, particularly when studies examine risks with small increments of exposure (*e.g.*, 10 µg/m³). Other studies estimated exposure to PM_{2.5} based on measurements of PM₁₀ exposures, and this could bias results in either direction (Beelen *et al.*, 2008; Laden *et al.*, 2000).

All of these factors make it difficult to attribute risks to long-term exposure to PM_{2.5}. More importantly, the long-term studies relied on in the ISA had few exposures below 15 µg/m³, so they were not informative regarding risks below the current NAAQS.

3.3 Section 2.4 does not accurately portray health impacts of PM_{2.5} exposure.

Section 2.4 of the ISA discusses the public health impacts associated with exposure to PM, including exposure-response relationships and evidence that certain populations are potentially susceptible or vulnerable to PM exposure. The epidemiology data relied on in the ISA to assess the

concentration-response relationship are not robust enough to determine whether a linear no-threshold model best describes the association between PM exposure and health effects.

3.3.1 Studies relied on in the ISA to assess concentration-response relationships are not sufficient for concluding a linear model.

For evaluating the concentration-response relationship between mortality and short-term exposure to PM, the ISA relied on studies by Daniels *et al.* (in HEI, 2004), Schwartz (2004), and Samoli *et al.* (2005). The ISA suggests that linear no-threshold models best describe the associations between PM and health effects. Evidence suggests that these may not be the best models to describe this association, however.

Daniels *et al.* (in HEI, 2004) analyzed three possible models to describe the relationship between PM₁₀ and mortality and concluded that a log-linear model was the most appropriate model for both cardiorespiratory and total mortality (but not other-cause mortality). The models were analyzed using Akaike Information Criterion (AIC), however, and as discussed here in Section 6.4, this may not be an appropriate criterion upon which to base the choice between models.

Schwartz (2004) used a different technique, including indicator variables for days in which the PM₁₀ concentration was within specific ranges, and did not find evidence for non-linearity when combining estimates across 14 cities. This study did not analyze city-specific thresholds, however, and heterogeneity in the concentration-response curve across cities was not examined. Heterogeneity across cities may influence the shape and gradient of the concentration-response relationship, and can make a non-linear relationship appear linear at low exposure levels.

Samoli *et al.* (2005) observed heterogeneity in the shape of the concentration-response curves across 22 European cities. Their analysis supported a log-linear association between PM₁₀ and mortality, but the ISA correctly stated that "the heterogeneity observed between cities complicates the biological explanation for the combined and city-specific results." The ISA further states: "Overall, the aforementioned studies all support the use of a no-threshold log-linear model, but additional issues such as the influence of heterogeneity in estimates between cities, and the effect of seasonal and regional differences in PM on the C-R relationship still require further investigation."

For mortality associated with long-term exposure to PM, the ISA relied on studies by Schwartz *et al.* (2008) and Roman *et al.* (2008). Schwartz *et al.* (2008) examined the concentration-response relationship between PM_{2.5} and mortality using data from the Harvard Six Cities Study. Two approaches were used, each involving Cox proportional hazards models, and both approaches found that the concentration-response curve was indistinguishable from linear. The use of the Cox model may have led to biased results (see Section 2.1.4).

Roman *et al.* (2008) developed probabilistic uncertainty distributions to characterize uncertainties in the concentration-response relationship for annual PM concentrations ranging from 4 to 30 µg/m³. A panel of 12 experts was asked to provide judgment on the true shape of the concentration-response curve. The majority of the panel agreed that, collectively, the epidemiologic data did not provide evidence of a population threshold. Several of these experts were authors of key air pollution studies, however, so they may have had preconceived opinions regarding the nature of the concentration-response relationship. Other underlying cognitive tendencies that influence expert judgment, but cannot be accounted for, include the tendency to assign greater probability to frequently mentioned events, the tendency to be over-influenced by the first pieces of information provided, and the tendency of experts to overestimate the probability that their opinions are correct. In addition, the study by Roman *et al.* (2008) emphasized the conclusions of the expert panel, but not the data that went into these conclusions. Thus, one cannot evaluate the validity of their analyses.

In the epidemiology studies relied on by the ISA, many of the uncertainties within the data (such as confounding, measurement error, and exposure misclassification) were not accounted for in the statistical models. Even if a linear model best describes the reported data, it is plausible that a non-linear model would have better described the data were these uncertainties taken into account. Thus, the currently-available PM epidemiology data are simply not robust enough to determine whether a linear no-threshold model best describes the association between PM exposure and health effects.

4 Studies of short-term PM_{2.5} exposure, such as those reviewed in Chapter 6 of the ISA, do not support a causal association between PM_{2.5} at concentrations below the current NAAQS and health effects.

In the ISA, US EPA concludes that the association between short-term exposure to PM_{2.5} and cardiovascular morbidity is "causal" and that the associations between short-term exposure to PM_{2.5} and respiratory morbidity and mortality are "likely to be causal." There have been no studies published since the 2004 AQCD that provide reduced uncertainties or stronger evidence for the previously identified effects; provide evidence that risks for the previously identified effects have increased since the last review; or provide further information on the possibility that these effects occur at lower levels than previously identified. Several major studies are described below that do *not* support effects at PM_{2.5} levels below the current NAAQS. Several of these studies reported either null or weakly positive findings. In other cases, weakly positive findings became non-significant when adjusted for confounders. Several studies did not have information on co-pollutants, so reported associations were likely biased away from the null. Risk estimates across cities were heterogeneous in many of these studies; thus, pooling the risk estimates together was inappropriate. Risk estimates in most of the studies were sensitive to the various lag times investigated, and statistically significant associations were reported for particular effects at different lag times across the studies. Exposure misclassification is perhaps the biggest shortcoming of the short-term studies of PM_{2.5}, as all of the studies reviewed here used measurements from central monitors as surrogates for personal exposures. All of these factors make it difficult to attribute risks to short-term exposure to PM_{2.5}, and several of these studies did not report actual exposures. Thus, these short-term studies are not informative regarding risks below the current NAAQS.

4.1 Short-term studies of cardiovascular and respiratory morbidity do not provide support for a causal association at levels below the current NAAQS.

In Section 6.2.11.3, the ISA concludes that epidemiology studies "provide support" for associations between PM_{2.5} and hospital admissions for cardiovascular diseases in areas with mean concentrations ranging from 13.8 to 18.8 µg/m³. In addition, in section 6.3.9.3, the ISA concludes that "Adverse associations between PM_{2.5} and hospitalizations and ED [emergency department] visits for respiratory diseases (*e.g.*, COPD and respiratory infections) have been consistently observed among older

adults while the associations of asthma hospitalizations and ED visits with PM_{2.5} are more heterogeneous." Below, two of the studies upon which US EPA based its conclusions are described (Dominici *et al.*, 2006; Bell *et al.*, 2008). Both studies assessed the relationship between short-term exposure to PM_{2.5} and hospital admissions for both cardiovascular and respiratory outcomes. US EPA should not consider these studies supportive of a causal association between PM_{2.5} and cardiovascular or respiratory morbidity at concentrations below the current NAAQS.

4.1.1 Dominici *et al.* (2006) based risk estimates on exposures measured with central monitors and Medicare data, did not account for co-pollutants, and found results to vary geographically.

In this multi-city Medicare Air Pollution Study (MCAPS), daily time-series data on hospital admission rates for cardiovascular and respiratory outcomes among Medicare beneficiaries (aged > 65 years) residing an average of 5.9 miles from a PM_{2.5} monitor were used to assess the associations between cause-specific hospitalization rates and same-day PM_{2.5} levels for 204 US urban counties from 1999-2002. The average of the county mean annual values of PM_{2.5} for this time period was 13.4 µg/m³ (interquartile range [IQR]: 11.3-15.2 µg/m³). This study used Poisson regression models controlling for long-term temporal trends and meteorologic conditions with natural cubic splines. County-specific results were averaged using Bayesian hierarchical models at 0, 1, and 2-day single lags, and three-day distributed lag models (lags 0, 1, and 2 days) were also considered in a subset of 90 counties with daily PM_{2.5} data available during the study period. The models combined relative risks across counties, accounting for within-county statistical error and for between-county variability of the "true" relative risks (as a test for heterogeneity).

The ISA states that Dominici *et al.* (2006) found positive associations between day-to-day variation in PM_{2.5} concentration and hospital admission for all outcomes for at least one exposure lag, but not all of these associations were statistically significant. For cardiovascular outcomes, statistically significant excess risks of 0.8%, 0.4%, and 1.3% per 10 µg/m³ increase in PM_{2.5} were reported for cerebrovascular disease (lag 0), ischemic heart disease (lag 2), and heart failure (lag 0), respectively. Excess risks were not statistically significant for peripheral vascular disease or outcomes related to heart rhythm for any lag time, although the ISA incorrectly stated that the excess risk for heart rhythm outcomes was significant at lag 0. Strong regional and seasonal heterogeneity was observed for cardiovascular outcomes with the strongest estimates in the Northeast. For respiratory outcomes, excess risks of 0.9% per 10 µg/m³ increase in PM_{2.5} were reported for both COPD (lag 0) and respiratory tract

infections (lag 2). Heterogeneity in effect estimates for respiratory outcomes was observed across the US with an association close to null reported in the Northeast.

For cardiovascular outcomes across the various lag times, some associations were negative or null, and most associations were not statistically significant. For respiratory outcomes, the associations were positive across all lag times but not all were statistically significant. The authors focused only on the lag times that were significant for each outcome.

This study has several limitations. One is related to the outcome data; Medicare data are collected for administrative purposes, and diagnoses can be subject to some degree of misclassification and can vary geographically. As with all ecological studies, individual data were not available, so it was not possible to associate specific exposures with specific outcomes. The exposures were measured at the county level, using central monitors, so they are purely ecological. This study did not control for effects of other potentially confounding air pollutants (*e.g.*, NO_x, SO_x, and O₃). Finally, it is not clear why there were statistically significant findings in the eastern, but not the western, US.

Although the average level of PM_{2.5} across counties in this study was below the current NAAQS, the lack of consistent, statistically significant results across lag times and the limitations described above suggest that this study should not be used to assess the association between PM_{2.5} and cardiovascular or respiratory health effects.

4.1.2 Bell *et al.*, (2008) based risk estimates on exposures measured with central monitors and Medicare data, did not account for co-pollutants, and found results to vary geographically.

Bell *et al.* (2008) extended the MCAPS database from the Dominici *et al.* (2006) study and investigated whether short-term effects of PM_{2.5} on risk of cardiovascular and respiratory hospitalizations among the elderly varied by region and season in 202 US counties from 1999-2005. The average PM_{2.5} level across these counties was not reported, but the authors stated that the IQR of overall PM_{2.5} levels was 8.7 µg/m³.

Three different first-stage models were used to estimate associations within single counties, then a second-stage model combined county-specific estimates, accounting for their statistical uncertainty to generate an overall effect. The first-stage main effect model assumed that the effect on hospitalizations is constant throughout the year. A second "seasonal" model allowed the effect to vary by season, and a

third "harmonic" model allowed the effect to vary smoothly throughout the year and was used as a sensitivity analysis. The main effect and seasonal interaction models were applied for admissions at lags 0, 1, and 2 days, and the harmonic model was applied for the lag with the strongest effects, as reported by the other two models. All three models were fitted separately within geographic regions. Regional analyses included 200 counties. Evidence of seasonal and regional heterogeneity was tested for the lags with the strongest effects for each hospitalization cause. The Wald test statistic was used to assess for evidence of heterogeneity in national average effects across seasons and regional average effects across regions for both cardiovascular and respiratory admissions. The Wald test statistic was compared with a chi-square distribution with appropriate degrees of freedom to obtain corresponding significance levels.

Bell *et al.* (2008) found evidence of substantial and statistically significant variability in the effects of PM_{2.5} on cardiovascular hospitalizations by season and region, with the highest effects in winter and in the Northeast. Nationwide excess risk of cardiovascular admissions was 0.8% per 10 µg/m³ increase in PM_{2.5} at lag 0 and was not statistically significant at other lag times. Across seasons and regions, most associations were not statistically significant at any lag time.

The ISA states that Bell *et al.* (2008) observed "largely null findings" for PM_{2.5} and respiratory hospitalizations (COPD, respiratory tract infections) for the US as a whole but reported heterogeneity in effect estimates across the country that were explained by regional and seasonal factors. Nationwide excess risk was only statistically significant at lag 2, and across seasons and regions, most associations were not statistically significant at any lag time. In contrast, later in the report, the ISA states that Bell *et al.* reported "consistent" associations of PM_{2.5} with COPD and respiratory infections, which is incorrect, particularly when only statistically significant associations are considered.

This study has similar limitations to those of the Dominici *et al.* (2006) study. Medicare data are collected for administrative purposes, and diagnoses are known to be subject to some degree of misclassification and to vary geographically. The exposures were measured at the county level, using central monitors, so they are purely ecological. As with all ecological studies, individual data were not available, so it was not possible to associate specific exposures with specific outcomes. In addition, this study did not control for effects of other potentially confounding air pollutants, which likely biased results away from the null. Finally, it is not clear why there were statistically significant findings in the eastern, but not the western, US.

4.2 Short-term studies of cardiovascular and respiratory mortality do not provide support for a causal association.

In section 6.5.3.2, the ISA concludes that "the epidemiologic evidence on the effect of short-term exposure to PM_{2.5} on mortality is sufficient to conclude that a causal relationship is likely to exist at ambient concentrations." Below, four of the major studies published since the 2004 AQCD that examined short-term exposure to PM_{2.5} and mortality are reviewed (Franklin *et al.*, 2007; Ostro *et al.*, 2006; Burnett *et al.*, 2004; Dominici *et al.*, 2007). These studies do *not* support effects at PM_{2.5} levels below the current NAAQS.

4.2.1 Franklin *et al.* (2007) calculated risk estimates that were heterogeneous (and higher in cities with lower PM_{2.5} levels), did not account for co-pollutants, and were sensitive to model selection.

Franklin *et al.* (2007) analyzed the relationship between exposure to PM_{2.5} and mortality in 27 US cities with PM_{2.5} monitoring and daily mortality data for at least two years of a 6-year period from 1997-2002. The mean concentration of PM_{2.5} across all cities was 15.7 µg/m³. Effect modification of age and gender were examined using a case-crossover model while effect modification by geographic location, annual PM_{2.5} concentration above and below 15 µg/m³, and use of central air conditioning were estimated using meta-regression. At lag 1 day, they reported a 1.2%, 0.94%, 1.8%, and 1.0% increase in all-cause mortality, cardiovascular mortality, respiratory mortality, and stroke deaths, respectively, per 10 µg/m³ of PM_{2.5}. They reported that statistically significant effect modification occurred by age and geography: the effects of exposure were greater in subjects ≥ 75 years of age, in eastern cities, and for those without central air conditioning. Franklin *et al.* (2007) also reported non-significant higher risk estimates in cities with annual PM_{2.5} concentrations below 15 µg/m³.

Based on the city-specific risk estimates, there was no clear association between PM_{2.5} and all-cause mortality. Risk estimates were negative for seven cities, with three of these being statistically significant. In five other cities, the risk estimates were near null. Negative or null associations were observed in many of the cities with the highest daily average concentrations of PM_{2.5}. The authors provided no basis for selecting the results at lag 1 as being definitive in this analysis. The lag 1 results showed an association for respiratory mortality, but not cardiovascular mortality, which is not consistent with results of chronic studies which demonstrate the opposite finding. The ISA notes that the wide

confidence intervals associated with the risk estimates for each effect modifier (*e.g.*, age, geographic location, air conditioning use, PM levels above or below 15 $\mu\text{g}/\text{m}^3$) suggest low statistical power for testing the differences between effect modifiers.

The test for heterogeneity was highly statistically significant for all-cause, cardiovascular, and respiratory mortality, and the risk estimates for these mortality indicators were unadjusted for factors that could account for the heterogeneity (*e.g.*, confounding by co-pollutants such as O_3 or NO_2); thus, the data should not have been pooled. This heterogeneity indicates that the pooled risk estimates may not be valid, and also may explain why many of the estimates for these indicators across three lag times are not statistically significant.

The authors reported that health effects may be observed below the NAAQS standard because larger risk estimates were observed in cities with average ambient $\text{PM}_{2.5}$ values below 15 $\mu\text{g}/\text{m}^3$ than for cities with an average above this value. None of these risk estimates was statistically significant, however, and overall estimates across cities were highly heterogeneous, such that no trend can be discerned. Also, as stated above, risk estimates for individual cities with some of the highest daily average concentrations of $\text{PM}_{2.5}$ were negative or null, implying a reverse dose-response that is not biologically plausible. The data from this study are insufficient to address whether there is an association between $\text{PM}_{2.5}$ at levels below the current NAAQS and mortality.

4.2.2 Ostro *et al.*, (2006) calculated risk estimates that appeared to be heterogeneous and were sensitive to co-pollutants and model selection.

Ostro *et al.* (2006) examined the association between $\text{PM}_{2.5}$ and daily mortality in nine heavily-populated California counties from 1999-2002. Mean daily $\text{PM}_{2.5}$ levels ranged from 14 to 29 $\mu\text{g}/\text{m}^3$ across counties. All-cause and several cause-specific subcategories of mortality were considered. The associations were examined among several subpopulations, including the elderly (> 65 years of age), males, females, non-high school graduates, whites, and Hispanics. Ostro *et al.* (2006) used Poisson multiple regression models incorporating natural or penalized splines to control for covariates that could affect daily counts of mortality, including time, seasonality, temperature, humidity, and day of the week. They used meta-analyses using random-effects models to pool the observations in the nine counties. A 10 $\mu\text{g}/\text{m}^3$ change in 2-day average (of 0- and 1-day lags) $\text{PM}_{2.5}$ concentration corresponded to increases of 0.6%, 0.6%, 2.2%, 2.4%, and 0.7% for all-cause, cardiovascular, respiratory, diabetes, and elderly mortality, respectively, in the pooled analysis using penalized splines.

Ostro *et al.* (2006) did not test for heterogeneity before combining the county-specific risk estimates, which appear to be heterogeneous. Thus, it is questionable whether these data should have been combined. The heterogeneity in this study cannot be explained by factors that are usually hypothesized to explain heterogeneity in time-series studies, such as seasonality or geographic differences. The ISA notes the wide confidence intervals associated with the risk estimates for each effect modifier in the combined analysis, suggesting low statistical power for testing the differences between effect modifiers.

The results of this study were sensitive to the lag time, model specification, and type of spline used. Risk estimates from using a single day lag of 2-days (lag 2) were lower for all-cause, cardiovascular, respiratory, and elderly mortality and were only statistically significant for respiratory mortality compared to those using the average of 0- and 1-day lags. With the natural spline model, the percent change in daily mortality decreased as the degrees of freedom/year increased, and almost all of the risk estimates were not statistically significant using this model. These risk estimates were also lower than those reported using the penalized spline model. It is not clear why Ostro *et al.* (2006) chose to focus on the results using penalized splines or how they selected what they considered to be the optimal degrees of freedom.

Ostro *et al.* (2006) stated that when NO₂ and CO, co-pollutants that are highly correlated with PM_{2.5}, were included in the model, they tended to attenuate the magnitude and significance of the PM_{2.5} coefficient. The actual results from the two-pollutant models were not presented, but the fact that the results shifted to become non-statistically significant suggests the possibility that the attenuation may have been rather high. Given such findings, it is unclear as to how the increased risks in mortality can be attributed solely to PM_{2.5}.

In contrast to most other studies assessing the relationship between PM_{2.5} and mortality, the increased risks for cardiovascular mortality in this study were lower (by three-fold) than those for respiratory mortality. Because of this and the other limitations discussed above, this study should not be used as evidence for an association between PM_{2.5} at levels below the current NAAQS and mortality.

4.2.3 Burnett *et al.* (2004) identified NO₂ as a confounder of PM_{2.5} risk estimates in time-series studies.

Burnett *et al.* (2004) examined the association between mortality and average daily variations in PM_{2.5} (sampled every sixth day), NO₂ (sampled daily), and other air pollutants in twelve large Canadian cities, using a 19-year time-series analysis from 1981-1999. The average concentration of PM_{2.5} across cities was 12.8 µg/m³. This study found that a 12.8 µg/m³ increase in PM_{2.5} at lag 1 was associated with a 0.77% increase in non-accidental mortality. After adjustment for NO₂, however, the risk estimate was negative. In addition, the authors examined the same associations using levels of both PM_{2.5} and NO₂ that were sampled daily in 11 of the 12 cities from 1998-2000. This analysis yielded a 1.13% increase in mortality per 10 µg/m³ increase in PM_{2.5}, but this risk estimate decreased to 0.98% upon adjustment for NO₂. These results strongly suggest that NO₂ is a confounder of the association between PM_{2.5} and mortality in time-series studies and should be considered in the evaluation of all ecological studies. This was noted in the ISA, and this study was not used as evidence for an association between PM_{2.5} and mortality.

4.2.4 Dominici *et al.* (2007) calculated risk estimates based on uncertain exposure estimates that were unadjusted for co-pollutants and were sensitive to model selection.

This study used data from the National Morbidity Mortality Air Pollution Study (NMMAPS) to estimate national average relative rates of the effects of PM_{2.5} on all-cause mortality across 96 counties from 1999-2000. The mean PM_{2.5} concentration was not provided for these counties; instead, the median county-specific concentrations for 567 and 682 US counties with available PM_{2.5} data in the years 1999 and 2000, respectively, were presented in box plots, and both were between 10 and 15 µg/m³. Bayesian two-stage hierarchical models were used to estimate county-specific, regional, and national average associations between day-to-day variation in PM_{2.5} at lag 1 and county-level mortality counts. City-specific estimates of risk associated with PM_{2.5} were pooled across the 96 counties, and heterogeneity was accounted for in the analysis. The models assumed linearity and were generated using smoothed time functions and 7 degrees of freedom per year to account for weather, seasonality, influenza epidemics, medical practice variation, and long-term trends in PM. The authors found that a 10 µg/m³ increase in PM_{2.5} at lag 1 was associated with 0.29% and 0.38% increases in all-cause and cardiorespiratory mortality, respectively, but the cardiorespiratory risk was not statistically significant.

The study by Dominici *et al.* (2007) suffers from several limitations. The authors noted the large degree of statistical uncertainty in the PM_{2.5} risk estimates because of the availability of only two years of PM_{2.5} data. They did not adjust for simultaneous exposure to co-pollutants, and the results were sensitive to lag and degrees of freedom. Risk estimates were statistically insignificant above 7 degrees of freedom. No justification was given for selection of the particular lag and degrees of freedom used in the study, so it is not clear which models were most appropriate. This study used data from monitors as a surrogate for personal exposure; thus, exposure misclassification may have biased the estimates in all of the models. PM_{2.5} concentrations for the counties used in this study were not provided, so it not known if any of the risk estimates are applicable at levels below the current NAAQS.

4.3 Selected studies investigating the effects of short-term exposure to PM_{2.5} do not support a causal association with morbidity or mortality at levels below the current NAAQS.

In the ISA, US EPA concludes that the association between short-term exposure to PM_{2.5} and cardiovascular morbidity is "causal" and that the associations between short-term exposure to PM_{2.5} and respiratory morbidity and mortality are "likely to be causal." I reviewed two studies upon which US EPA based its conclusions regarding short-term PM_{2.5} exposure and cardiovascular and respiratory morbidity (Dominici *et al.*, 2006; Bell *et al.*, 2008) and four studies that examined short-term PM_{2.5} exposure and mortality (Franklin *et al.*, 2007; Ostro *et al.*, 2006; Burnett *et al.*, 2004; Dominici *et al.*, 2007). Based on the review of these short-term exposure studies, I conclude that the new epidemiology studies do not support a causal association between PM_{2.5} at concentrations below the current NAAQS and health effects. That is, these studies should not be used as evidence of a "causal" or "likely causal" relationship between short-term exposure to ambient PM_{2.5} and cardiovascular morbidity, respiratory morbidity, or mortality.

Several of these studies reported either null or weakly positive findings (Dominici *et al.*, 2006; Bell *et al.*, 2008, Franklin *et al.*, 2007). In other cases, weakly positive findings became non-significant when adjusted for confounders (Ostro *et al.*, 2006; Burnett *et al.*, 2004). Several studies did not have information on co-pollutants, so reported associations were likely biased away from the null (Dominici *et al.*, 2006, 2007; Bell *et al.*, 2008, Franklin *et al.*, 2007). Risk estimates across cities were heterogeneous in many of these studies (Dominici *et al.*, 2006; Bell *et al.*, 2008, Franklin *et al.*, 2007) so pooling cities was inappropriate. Risk estimates in most of the studies were sensitive to the various lag times

investigated (Dominici *et al.*, 2006, 2007; Bell *et al.*, 2008, Franklin *et al.*, 2007; Ostro *et al.*, 2006). Moreover, statistically significant associations were reported for particular effects at different lag times across the studies. For example, Dominici *et al.* (2006) reported a statistically significant increased risk for COPD at lag 0, but Bell *et al.* (2008) reported a statistically significant increased risk for this same outcome only at lag 2.

Exposure misclassification/measurement error is perhaps the biggest shortcoming of the short-term studies of PM_{2.5}. All of the studies reviewed here used measurements from central monitors as surrogates for personal exposure. The distance between people's residences to these monitors can vary, increasing the likelihood of obtaining inaccurate measurements.

All of these factors make it difficult to attribute risks to short-term exposure to PM_{2.5}, and several of these studies did not report actual exposures (Bell *et al.*, 2008; Dominici *et al.*, 2007). Thus, these short-term studies are not informative regarding risks below the current NAAQS.

5 Studies of long-term PM_{2.5} exposure, such as those reviewed in Chapter 7 of the ISA, do not support a causal association between PM_{2.5} at concentrations below the current NAAQS and health effects.

In the ISA, US EPA concludes that the associations between long-term exposure to PM_{2.5} and cardiovascular morbidity, respiratory morbidity, and mortality are "likely to be causal" and that the association between long-term exposure to PM_{2.5} and reproductive and developmental outcomes is "suggestive." Based on an analysis of the recent epidemiology literature, I conclude that there have been no studies published since the 2004 AQCD that suggest uncertainties have been reduced; or that these associations provide stronger evidence for the previously identified effects; provide evidence that risks for the previously identified effects are higher than previously estimated; or provide further information on the possibility that these effects occur at lower levels than previously identified. Several major studies are described below, some of which were published after the first external review draft of the PM ISA was released, that do *not* support long-term PM_{2.5} exposures leading to these health effects at levels below the current NAAQS. Several of these studies reported either null or weakly positive findings. In other cases, weakly positive findings became non-significant when adjusted for confounders. Several studies did not have information on co-pollutants or other factors that may have been associated with exposure and/or outcome, so reported associations were likely biased away from the null. Exposure misclassification is perhaps the biggest shortcoming of the long-term studies of PM_{2.5}, as all studies reviewed here used measurements from central monitors as surrogates for personal exposures. All of these factors make it difficult to attribute risks to long-term exposure to PM_{2.5}. More importantly, the long-term studies relied on in the ISA had few exposures below 15 µg/m³, so they were not informative regarding risks below the current NAAQS.

5.1 Long-term studies of cardiovascular morbidity do not provide support for a causal association.

In Section 7.2.7.3, the ISA concludes: "Epidemiologic evidence of the adverse effect of PM_{2.5} on subclinical markers of atherosclerosis is available from the majority of recent studies on this topic. In addition, a large US study reports associations of 1-year average PM_{2.5} concentration with cardiovascular diseases among post-menopausal women. Further, modification of the PM_{2.5}-CVD association by

smoking status and use of anti-hyperlipidemics has been reported in more than one epidemiologic study." Below, four studies upon which US EPA based its conclusions are described, two of which reported subclinical effects (Allen *et al.*, 2009; Diez Roux *et al.*, 2008), and two of which describe clinical outcomes in epidemiological studies (Hoffman *et al.*, 2006; Miller *et al.*, 2007). US EPA should not consider these studies supportive of a causal association between PM_{2.5} and cardiovascular morbidity at concentrations below the current NAAQS.

5.1.1 Allen *et al.* (2009) did not report statistically significant risk estimates for subclinical measures of atherosclerosis.

The ISA cites the study by Allen *et al.* (2009) as supporting an association between subclinical measures of atherosclerosis and long-term PM_{2.5} exposures, with larger increases among users than non-users of anti-hyperlipidemics. Allen *et al.* (2009) conducted a cross-sectional analysis of exposure to PM_{2.5}, residential proximity to major roadways, and the presence and extent of abdominal aortic calcification. Using 1,147 randomly selected subjects from five cities enrolled in the Multi-Ethnic Study of Atherosclerosis (MESA) Aortic Calcium Ancillary Study, the authors gathered data including residential information, SES variables, and potential risk factors. Each subject was scanned for abdominal aortic calcification using computed tomography. The authors assigned PM_{2.5} values based on the average concentrations over a two-year period (2000-2002) obtained from US EPA's Aerometric Information Retrieval Service (AIRS, which is now the AQS database). The authors conducted a two-part analysis based on the presence of calcification and then, for those with calcification, the extent of calcification (using the Agatston score, which is a measurement based on which is based on the area and the density of calcified plaques). The authors adjusted analyses for some confounders and tested for effect modifiers.

The authors reported "a slightly elevated risk of aortic calcification (RR = 1.06; 95% CI: 0.96-1.16) with a 10 µg/m³ increase in PM_{2.5}." Because this is not statistically significant, it should not be considered "slightly elevated." Although PM_{2.5} exposure alone had no effect on the extent of calcification, Allen *et al.* (2009) reported that PM_{2.5} effect on the Agatston score was "significant" ($p_{interaction} = 0.06$) for users of lipid-lowering medications. They noted that the "interpretation of this finding in a cross-sectional analysis is complicated by the fact that duration of the medication use was not considered, and lipid-lowering medications slow the progression of abdominal aortic plaques, and may therefore, reduce progression of calcification."

Using recent measures of PM_{2.5} for assessing the effect of exposure on the development of a chronic disease is a re-occurring issue in the epidemiology literature. The authors acknowledged:

Our exposure assessment approach relied on the strong assumption that the 2-year average PM_{2.5} was representative of longer-term past exposures. We were able to assess the relationship only between calcification and this relatively recent exposure information, even though the development of calcification is a long-term process that may be affected by air pollution exposures over the full lifetime.

Setting aside these study limitations, the average PM_{2.5} for all five cities was $15.8 \pm 3.6 \mu\text{g}/\text{m}^3$, which is comparable to the current annual standard for PM_{2.5} (no analyses were presented for individual cities). All analyses were conducted using subjects from all cities, meaning none were based on average PM_{2.5} levels below the current standard.

5.1.2 Diez Roux *et al.* (2008) reported mostly null associations (based on linear models) between PM_{2.5} and subclinical measures of atherosclerosis.

The ISA states that Diez Roux *et al.* (2008) reported associations between long-term PM_{2.5} exposure and subclinical measures of atherosclerosis. These investigators evaluated approximately 5,000 US adults as part of the MESA. They specifically examined associations between long-term exposure to PM_{2.5} and coronary calcium (CAC), common carotid intimal-medial thickness (CIMT), and ankle-brachial thickness (ABI) in adults without cardiovascular disease. Extensive information on demographics, CVD risk factors, and SES were obtained from personal interviews conducted between 2000 and 2002. The past 20-year individual PM_{2.5} exposure was estimated based on a self-reported residential history, which included any move dates, starting from 1982. Monthly mean PM_{2.5} measures were calculated from EPA community air monitor data. A spatio-temporal model was used to predict PM_{2.5} exposures based on the geocoded location of each participant's residence relative to central monitors for each participant month. For PM_{2.5} levels, the authors used two different metrics – actual measurements taken in 2001 (mean $16.7 \mu\text{g}/\text{m}^3$; SD, $3.8 \mu\text{g}/\text{m}^3$) and a 20-year reconstruction model (mean $21.7 \mu\text{g}/\text{m}^3$; SD $7.5 \mu\text{g}/\text{m}^3$) – which were highly correlated ($r = 0.64$). Of the six MESA center sites, only Minnesota had a mean PM_{2.5} level below $15 \mu\text{g}/\text{m}^3$. Diez Roux *et al.* (2008) reported that there was no effect of the study site location on the outcome measures.

Diez Roux *et al.* (2008) examined associations between the difference in PM_{2.5} concentration between the 10th to the 90th percentiles ($12.5 \mu\text{g}/\text{m}^3$) and subclinical effects. A total of 16 analyses (4 outcomes: CIMT, ABI, CAC, and extent of CAC; 2 measures of PM_{2.5}: input model, 2001 actual mean;

and 2 levels of model adjustment: demographic and SES factors, demographic, SES, and risk factors) revealed only one outcome (CIMT) that was weakly positive, the rest were null. The maximum effect was seen in the PM_{2.5} 2001 mean with full adjustments (Relative difference =1.03; 95% CI: 1.01-1.05). There was no effect of exposure on either ABI, presence of CAC, and if present, the extent of CAC. The authors also found, "no evidence that long-term particulate matter exposure was more strongly associated with subclinical disease in subgroups previously hypothesized to be more vulnerable to these effects, including women, older persons, persons with hyperlipidemia or diabetics, obese persons, and persons whose educational levels are low."

The major strength of this study is the authors' extensive efforts to better characterize past exposures with their innovative temporal-spatial modeling of PM exposures. Still, this study is an ecologic study without individual measures of total PM_{2.5} exposure profiles. The authors acknowledged that central outdoor monitors might be "poor proxies for personal exposures." Although data on co-pollutant levels were collected, they were used only to develop the spatio-temporal exposure model and were not included in any analysis to assess confounding or effect modification. Any bias that may have occurred with self-reporting is probably as great, or greater, than the effect measures. Therefore, any of the weak associations reported for CIMT could result from bias/imprecision in the confounding alone. The final model specification was linear no-threshold, as the Generalized Additive Model did not show "clear evidence" of threshold effects but, as discussed in Section 6.4 of the PM ISA and Section 6 here, it is possible that if the uncertainties discussed above had been accounted for, the concentration-response may have been non-linear. Owing to overall null findings and the reliance on a linear model for a non-cancer subclinical outcome, this study is not suitable for setting a standard.

5.1.3 Hoffman *et al.* (2006) did not find a sizable influence of background PM_{2.5} on CHD morbidity.

With regard to studies of clinical outcomes of cardiovascular disease in epidemiology studies, the ISA discusses studies by Miller *et al.* (2007) (discussed here in Section 5.3.2) and Hoffman *et al.* (2006):

Two epidemiologic studies of the PM_{2.5}-CVD morbidity relationship focused on clinical CVD events: in one case, on incident, validated MI, coronary revascularization, and stroke in 36 U.S. metropolitan areas (Miller *et al.*, 2007b), and in the other, on prevalent, self-reported CHD in Essen and Mülheim, Germany (Hoffmann *et al.*, 2006). Miller *et al.* (2007b) was a prospective, cohort study with the population restricted to women (Miller *et al.*, 2007b). Authors used arithmetic averaging of year 2000 AQS PM_{2.5} data at the monitor most proximate to each participant's geocoded U.S. Postal Service ZIP code.

The one year average PM_{2.5} exposure used in the German study was based on dispersion-modeled emissions data (Hoffmann et al., 2006). The inconsistent findings between these two studies may be driven by differences in study design and location. Miller et al. (2007b) found large increases in the adjusted risk of MI, revascularization, and stroke with standardized increments in PM_{2.5}, but for the same increment, Hoffman et al. (2006) found no such increase in the odds of prevalent CHD. Furthermore, striking evidence for effect modification by anthropometric measures (e.g., BMI and waist-to-hip ratio) presented by Miller et al. (2007b) has been tested, e.g. by Diez Roux et al. (2008), but not observed again within this body of literature.

Hoffman *et al.* (2006) evaluated the effect of traffic exposure on CHD in 3,399 participants of the German Heinz Nixdorf RECALL study. High personal traffic exposure was defined as residing within 150 m of at least one major road, and CHD was defined as a self-reported history of MI, implantation of a coronary stent, angioplasty, or bypass surgery. Evaluations were collected from December 2000 until July 2003. Regional PM_{2.5} background levels were estimated from yearly mean values run through EURAD for the year 2002. For the whole population, the authors reported a significant effect of traffic exposure on CHD (OR = 1.62; 95% CI: 1.12-2.34), which rose to 1.85 (95% CI: 1.21-2.84) after adjusting for "background" PM_{2.5} and cardiovascular risk factors; adjustment for PM_{2.5} alone had no effect. The risks associated with PM_{2.5} in the four adjusted models were null (ORs ranged from 0.55 to 0.92 and the bounds of the 95% CIs ranged from 0.14 to 2.39).

The authors concluded that their "study demonstrates an association between the long-term residential exposure to traffic and prevalence of CHD." They also noted, "As can be expected from the small PM_{2.5} exposure contrast, we were not able to demonstrate a sizable influence of background PM_{2.5} on CHD morbidity." In fact, Table 1 shows that there is no difference in the mean and variability of PM_{2.5} between the low-traffic exposed (23.3 µg/m³, 1.4) and the high-traffic exposed (23.4 µg/m³, 1.4) groups. None of the subjects experienced an estimated exposure of background PM_{2.5} < 20 µg/m³, although it should be noted that PM_{2.5} levels are estimates, are considered "background levels," and are not measures of individual exposure.

The major strength of this study is the comprehensive assessment of individual-level information on major causal and conditional CHD risk factors, thus controlling for numerous confounders. Weaknesses of this study relate to reliable exposure assessment for the participants. Assumptions included that the current home address, thus the traffic exposure, is similar to past exposures during and/or preceding the CHD event. The authors were unable to account for instability in exposure estimates, which are based on a single year. In addition, the authors did not have information on the duration of residence. They noted that the number of relocations was small (<1% per year for the

evaluation period), but the health event occurred any time over a period of several years, and it is not clear whether the CHD event occurred while they were residing at their current address. Furthermore, the mobility between high-exposed and low-exposed individuals could have differed, and this was not considered. Hoffman *et al.* (2006) also acknowledged that the absences of information on occupational exposures and indoor exposures are limitations to the study. They did not, however, discuss how other traffic-derived air pollutants, such as CO, NO_x, or VOCs, could have biased their results. As alluded to in the ISA, this study does not provide evidence for cardiovascular morbidity at exposure concentrations below the current NAAQS annual level of 15 µg/m³.

5.1.4 Miller *et al.* (2007) jointly assessed morbidity and mortality, so effects of PM_{2.5} on morbidity alone could not be determined.

Morbidity analyses in the Miller *et al.* (2007) study are based on nonfatal and fatal events combined. Because morbidity and mortality were assessed jointly, this study cannot be used to draw conclusions between PM_{2.5} and morbidity. A discussion of this study is presented below, in Section 5.3.2.

5.2 Long-term studies of respiratory morbidity do not support a causal association.

Similar to its conclusion with regard to long-term exposure to PM_{2.5} and cardiovascular morbidity, the ISA concludes that "the evidence is sufficient to conclude that the relationship between long-term PM_{2.5} exposure and respiratory morbidity is likely to be causal." One study US EPA relies on for this conclusion is that by Goss *et al.* (2004), described below. Although statistically significant associations were noted, the likelihood that these were due to other pollutants or other factors cannot be ruled out.

5.2.1 Goss *et al.* (2004) based risk estimates on exposures measured with central monitors and did not account for several confounders.

Goss *et al.* (2004) investigated the relationship between criteria air pollutants and exacerbations of cystic fibrosis (CF) in a cohort of 11,484 patients enrolled in the CF Foundation National Patient Registry in 1999 and 2000. Exacerbations were defined as experiencing a CF-related pulmonary event requiring hospital admission or home-use of IV antibiotics. Patient populations were characterized by age, gender, race, percentage change in FEV₁, body weight, airway colonization of *pseudomonas*

aeruginosa or *burkholderia cepacia*, pancreatic function, and the $\Delta F508$ gene mutation. Type of insurance coverage and median household income were used as surrogates for SES. Air pollution data was collected from AIRS and matched to the patients' zip codes. The overall mean concentration of $PM_{2.5}$ was $13.7 \pm 4.2 \mu g/m^3$.

After adjusting for gender, age, weight, race, air colonization, pancreatic dysfunction, and insurance status, Goss *et al.* (2004) observed an effect of $PM_{2.5}$, PM_{10} , and ozone on the number of exacerbations. The authors reported that in a single pollutant model, a $10 \mu g/m^3$ increase in $PM_{2.5}$ was associated with an increased risk of having two or more exacerbations (OR= 1.21; 95% CI: 1.07-1.33). When the authors included baseline percent predicted FEV_1 in the statistical model, the effect of $PM_{2.5}$ exposure was no longer statistically significant. When they analyzed the association between a $10 \mu g/m^3$ change in $PM_{2.5}$ and having one exacerbation (*vs.* none), the OR was 0.70 (95% CI: 0.59-0.98); comparison of two exacerbations with no exacerbation remained insignificant (1.13; 95% CI: 0.99-1.29).

The investigators discussed several shortcomings in their study. They suggested that the attenuation of pollutant effects by adjusting for lung function could be that lung function decline "may be intimately associated with chronic exposure to air pollutants and may be part of the causal pathways in worsening prognosis." The patients who experienced two or more exacerbations were clearly sicker, and despite adjusting for some of these parameters (including lung function), many relevant variables remained. The authors noted that, "Residual confounding caused by unmeasured risk factors is of concern." They discussed the absence of information on tobacco use, ETS exposure, spatial effects such as climate, weather patterns, and regional variation in medical practices. They did not report on the use of oral antibiotics, pharmacological agents to increase mucus clearance or control airway constriction, or anti-inflammatory agents.

This is an ecologic study for which there are no individual measures of personal exposure to $PM_{2.5}$. It lacks assessment of many potential confounders and variables relating to misclassification of exposure. Although the authors adjusted for numerous differences between the two study groups, some factors (such as education in home health care management, regular clearance of lung secretions, medication use) that may have contributed to differences in overall health status at the initiation of the study, remained unadjusted for in the analyses.

5.3 Long-term studies of mortality do not support a causal association.

In section 7.6.8, the ISA states: "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 in areas with mean concentrations from 14 to 29 µg/m³ (Figure 7-8)." In addition, regarding long-term exposure to PM_{2.5}, the ISA states:

Collectively, the evidence is sufficient to conclude that the relationship between long-term PM_{2.5} exposures and mortality is likely to be causal. When looking at the cause of death, the strongest evidence comes from mortality due to cardiovascular disease, with additional evidence supporting an association between PM_{2.5} and lung cancer mortality (Figure 7-8). There is little new evidence that supports an association between PM_{2.5} exposure and respiratory mortality (Figure 7-8), though the existing evidence from the Harvard Six Cities and ACS studies show a strong relationship with cardiopulmonary mortality (Figure 7-7).

Below, seven of the major studies published since the 2004 AQCD that examined PM_{2.5} exposure and mortality are critically reviewed (Beelen *et al.*, 2008; Eftim *et al.*, 2008; Jerrett *et al.*, 2005; Laden *et al.*, 2006; Miller *et al.*, 2007; Pope *et al.*, 2009; Zeger *et al.*, 2008). These studies do *not* support effects on mortality at PM_{2.5} levels below the current NAAQS.

5.3.1 Beelen *et al.* (2008) reported no association between PM_{2.5} and mortality.

The ISA stated that the results of the Beelen *et al.* (2008) study "add to the evidence that long-term exposure to traffic-related particulate ambient air pollution is associated with increased mortality." Beelen *et al.* (2008) used data from an ongoing Dutch cohort study (NLS-AIR Study) of 120,850 subjects who were followed from 1987 to 1996. Participants' home addresses in 1986 were correlated with ambient levels of PM_{2.5} extrapolated from PM₁₀ measurements collected during 1992-1996, SO₂, BS, NO₂, and measures of traffic intensity. The authors assessed exposure effects for the full cohort (FC) and a smaller case-control cohort (CC). Although there were fewer study subjects in the CC, more information was available regarding potential confounders. The ISA does not discuss any of the findings from the CC, however.

PM_{2.5} mean concentrations were estimated as 28.3 µg/m³ (SD: 2.1 µg/m³). Relative risks were estimated using the Cox proportional hazards model (limitations of this model were discussed above, in Section 2.1.4) and based on the difference between the 5th and 95th percentiles (10 µg/m³). Figure 1 in this study indicates that none of the estimated values for PM_{2.5} were below the current NAAQS annual

value of $15 \mu\text{g}/\text{m}^3$ (minimum value, $23.0 \mu\text{g}/\text{m}^3$). For both cohorts, there were no statistically significant effects of $\text{PM}_{2.5}$ exposure on deaths from natural causes (FC: RR = 1.06, 95% CI: 0.97-1.16; CC: RR = 0.86, 95% CI: 0.66-1.13), cardiovascular disease (FC: RR = 1.04, 95% CI: 0.90-1.21; CC: RR = 0.83, 95% CI: 0.60-1.15), respiratory disease (FC: RR = 1.07, 95% CI: 0.75-1.52; CC: RR = 1.02, 95% CI: 0.56-1.88), lung cancer (FC: RR = 1.06, 95% CI: 0.82-1.38; CC: RR = 0.87, 95% CI: 0.52-1.47), or other causes (FC: RR = 1.08, 95% CI: 0.96-1.23; CC: RR = 0.85, 95% CI: 0.65-1.12). What is notable is that, even at high levels of estimated exposure, there were no increases in risk for any type mortality based on analyses of both the FC and the CC.

5.3.2 Miller *et al.* (2007) reported high risks (vs. other cohorts) for mortality based on the Cox model, and did not account for co-pollutants or exposure misclassification.

Regarding the Women's Health Initiative (WHI) study by Miller *et al.* (2007), the ISA concludes:

The WHI study not only confirms the Six City Study and ACS Study associations with mortality in yet another well characterized cohort with detailed individual-level information, it also has been able to consider the individual medical records of the thousands of WHI subjects over the period of the study. This has allowed the researchers to examine not only mortality, but also related morbidity in the form of heart problems (cardiovascular events) experienced by the subjects during the study. As reported in this paper, this examination confirmed that there is an increased risk of cardiovascular morbidity, as well (see section 7.2.1). These morbidity co-associations with $\text{PM}_{2.5}$ in the same population lend even greater support to the biological plausibility of the air pollution-mortality associations found in this study.

Miller *et al.* (2007) studied 65,893 postmenopausal women enrolled in the Women's Health Initiative (WHI) study without previous cardiovascular disease in 36 US metropolitan areas from 1994 to 1998, with a median follow-up of 6 years. Hazard ratios (HR) for fatal and non-fatal cardiovascular disease were correlated with exposure to $\text{PM}_{2.5}$. The annual average concentration of $\text{PM}_{2.5}$ (from US EPA AIRS for the year 2000) was used as the exposure metric, and each woman was linked to one of 573 monitors. Levels of $\text{PM}_{2.5}$ ranged from 3.4 (Honolulu) to $28.3 \mu\text{g}/\text{m}^3$ (Riverside), with a mean individual exposure of $13.5 \pm 3.3 \mu\text{g}/\text{m}^3$.

For each increase of $10 \mu\text{g}/\text{m}^3$ of $\text{PM}_{2.5}$, the authors reported significant increases in risk for any fatal or non-fatal cardiovascular event (HR = 1.24; 95% CI 1.09-1.41), CHD (HR = 1.21, 95% CI: 1.04-1.42); and stroke (HR = 1.28, 95% CI: 1.02-1.61), but not for MI (HR = 1.06, 95% CI: 0.85-1.34) or coronary revascularization (HR = 1.2, 95% CI: 1.00-1.43). They also reported an increased risk for

cardiovascular disease mortality (HR = 1.76; 95% CI, 1.25-2.47) associated with every 10 $\mu\text{g}/\text{m}^3$ increase in $\text{PM}_{2.5}$ exposure.

The strengths of this study include the extensive effort to characterize and confirm the presence of cardiovascular events. The authors also attempted to adjust for a wide range of factors that could potentially confound the interpretation of their findings.

Major drawbacks of this study include a lack of adjustment for co-pollutants and exposure misclassification. Although the majority of the subjects lived within 6 miles to their linked monitors, some lived as far as 30 miles away. Miller *et al.* (2007) also conducted their analyses using Cox proportional hazards models. It is not clear whether this was appropriate, as it is possible that the proportionality of hazards, either of $\text{PM}_{2.5}$ or of confounders, did not hold, leading to biased risk estimates (see Section 2.1.4). This seems like a plausible explanation when one considers that the risk estimates in this study were considerably higher than those in other cohorts, such as the Six Cities or ACS cohorts.

Another plausible explanation (not discussed in the ISA) is that Miller *et al.* (2007) overstated the risk of cardiovascular disease mortality by basing it on an incremental change of $\text{PM}_{2.5}$ of 10 $\mu\text{g}/\text{m}^3$, which is not attainable for most American cities (Jerrett and Burnett, 2007). Jerrett and Burnett (2007) note that the differences in $\text{PM}_{2.5}$ across cities is likely due to differences in sulfate PM levels, whereas the variation within each city is mostly driven by differences in $\text{PM}_{2.5}$ from traffic. Because these two components of $\text{PM}_{2.5}$ have different toxicities, the exposure increment Miller *et al.* (2007) used to interpret their hazard ratio should have reflected this difference. Jerrett and Burnett (2007) suggest that, when this difference is accounted for, the hazard ratio for cardiovascular disease mortality in New York, for example, decreases from 2.28 to 1.31. This is consistent with prior research.

Based on the high likelihood that risks are biased away from the null in this study, it should not be used to assess associations at levels below the current NAAQS.

5.3.3 Jerrett *et al.* (2005) calculated risks that were sensitive to model selection and were based on misclassified exposures and on Cox models that did not account for co-pollutants.

The ISA states that the results of the Los Angeles subset of the 1982-2000 ACS cohort (previously analyzed by Pope *et al.*, 2002) "suggest that previous and current studies may have

underestimated the magnitude of the association (Jerrett et al., 2005b)." Jerrett *et al.* (2005) estimated individual-level exposures by using several methods to spatially interpolate concentrations (based on data from the year 2000) from 25 PM_{2.5} monitors to residences of cohort members to 267 zip code areas in Los Angeles. The authors also assessed impact of traffic by proximity to freeways (within 500 or 1,000 meters). The set of 44 confounding variables used in previous ACS analyses were included as well as an additional eight ecologic variables relating to the neighborhood (*e.g.*, poverty, crime rate, racial composition, education, unemployment). Information on cause of death from 1982 to 2000 was available for 5,856 people in the Los Angeles subset.

Jerrett *et al.* (2005) used Cox proportional hazards models for their main analyses, and this may have biased findings away from the null (see Section 2.1.4). Although these investigators reported some statistically significant findings for risks of all-cause, IHD, cardiopulmonary, and lung cancer mortality, including some larger than those reported by Pope *et al.* (2002), findings for these same health outcomes were not statistically significant in other models with different sets of covariates. One cannot know with certainty which of these models is most appropriate. In addition, the only other pollutant adjusted for (and only in some models) was O₃; the potential effects of CO, SO_x, NO_x and other pollutants were not addressed. The lack of adjustment for these factors likely biased results away from the null. Exposure estimates also likely biased results away from the null, as exposures were based on PM_{2.5} levels in 2000, which were much lower than levels at the beginning of enrollment and before (when the disease processes leading to mortality began). It is also unlikely that exposures decreased uniformly throughout Los Angeles over the study period, which creates even more uncertainty in the analyses which assume the relative concentrations of PM_{2.5} are the same over the two-decade follow-up period. Risks of similar and greater magnitude in the same statistical models were reported for mortality from other causes which are not thought to be associated with PM_{2.5}, (*e.g.*, endocrine and digestive causes) suggesting the associations with IHD and cardiopulmonary mortality may be due to chance. These risks were not discussed in the ISA. Also, very few of the 267 zip codes are located in areas where interpolated PM_{2.5} concentrations were less than 14.4 µg/m³. Because this study is primarily based on PM_{2.5} levels above the current NAAQS, and is not informative regarding exposures below these levels.

5.3.4 Laden *et al.* (2006) calculated risk estimates based on central monitors and used Cox models that didn't account for co-pollutants or other confounders that changed over time.

The ISA relies heavily on data from the Six Cities Cohort. Regarding the recent study by Laden *et al.* (2006), it states:

A follow-up study has used updated air pollution and mortality data; an additional 1,368 deaths occurred during the follow-up period (1990-1998) vs. 1,364 deaths in the original study period (1974-1989) (Laden et al., 2006). Statistically significant associations are reported between long-term exposure to PM_{2.5} and mortality for data for the two periods (RR = 1.16 [95% CI: 1.07-1.26] per 10 µg/m³ PM_{2.5}). Of note, however, is a statistically significant **reduction** in mortality risk reported with **reduced** long-term fine particle concentrations (RR = 0.73 [95% CI: 0.57-0.95] per 10 µg/m³ PM_{2.5}). This is equivalent to an RR of 1.27 for reduced mortality risks. This reduced mortality risk was observed for deaths due to cardiovascular and respiratory causes, but not for lung cancer deaths. The PM_{2.5} concentrations for recent years were estimated from visibility data, which introduces some uncertainty in the interpretation of the results from this study. Coupled with the results of the original analysis (Dockery et al., 1993), this study strongly suggests that a reduction in fine PM pollution yields positive health benefits.

One of the major shortcomings of this analysis is that from 1979 to 1987, PM_{2.5} exposure data were measured directly from centrally located air-monitoring stations in each city, but after 1988, PM_{2.5} exposure was extrapolated from PM₁₀ AIRS data, humidity-corrected visibility data from local airports, and season indicators. Thus, it was not appropriate to conduct analyses using data from both time periods, as comparing data based on different metrics of exposure can lead to biased results, although it is not evident in which direction. Also, as shown in Figure 2 of the study, reproduced below, the association between PM_{2.5} and total mortality appears linear for Period 1 down to the lowest exposure estimate, but non-linear for Period 2 (lines added for emphasis). More importantly, for both periods, the lowest average PM_{2.5} concentration is > 10 µg/m³ and there is no information regarding how precise this estimate is (particularly considering uncertainties in exposure measurements). Thus, one cannot determine what the relationship between PM_{2.5} and all-cause mortality is around or below this point.

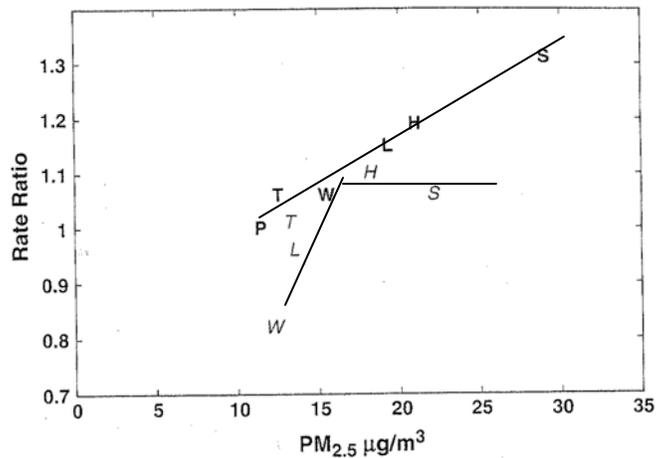


Figure 2. Estimated adjusted rate ratios for total mortality and PM_{2.5} levels in the Six Cities Study by period. P denotes Portage, WI (reference for both periods); T = Topeka, KS; W = Watertown, MA; L = St. Louis, MO; H = Harriman, TN; S = Steubenville, OH. A term for Period 1 (1 if Period 2, 0 if Period 1) was included in the model. *Bold letters* represent Period 1 (1974–1989) and *italicized letters* represent Period 2 (1990–1998). In Period 1, PM_{2.5} (µg/m³) is defined as the mean concentration during 1980–1985, the years where there are monitoring data for all cities (18). In Period 2, PM_{2.5} is defined as the mean concentrations of the estimated PM_{2.5} in 1990–1998.

Figure 5.1. Estimated adjusted rate ratios for total mortality and PM_{2.5} levels in the Six Cities Study by period. (Laden *et al.*, 2006, Figure 2, Lines added by Gradient)

Laden *et al.* (2006) estimated mortality rate ratios using Cox proportional hazards models. As discussed in Section 2.1.4, it is not clear whether this was appropriate, as it is possible that the proportionality of hazards, either of PM_{2.5} or of confounders, did not hold, leading to biased risk estimates. Also none of these models adjusted for co-pollutants, and this could have led to bias away from the null. Other factors entered into the models as potential confounding were based on measurements that occurred up to decades after their documentation, and one cannot know if they are still accurate. The opportunity for an unmeasured and influential confounder to have affected the RRs is also considerable in this study. Such residual confounding could be as strong as the effect estimates, which are weak, as is typical for air pollution epidemiological studies. Also, individual level covariates were not available in the Period 2 follow-up.

Although this study may be informative regarding exposures above the current annual NAAQS for PM_{2.5}, it does not provide substantially new information regarding exposures at levels below the current NAAQS.

5.3.5 Eftim *et al.* (2008) based risk estimates on exposures measured with central monitors and Medicare data and did not account for co-pollutants and other confounders.

Regarding long-term PM_{2.5} exposure studies in general, the ISA states:

Most recently, an ecological cohort study of the nation's Medicare population has also been completed (Eftim *et al.*, 2008). These new findings further strengthen the evidence linking long-term exposure to PM_{2.5} and mortality, while providing indications that the magnitude of the PM_{2.5}-mortality association is larger than previously estimated (Figure 7-8).

The ISA further states:

Using Medicare data, Eftim and co-authors (2008) have assessed the association of PM_{2.5} with mortality for the same locations included in the Six City Study and the ACS studies. For these locations, they estimated the chronic effects of PM_{2.5} on mortality for the period 2000-2002 using mortality data for cohorts of Medicare participants and average PM_{2.5} levels from monitors in the same counties included in the two studies. Using aggregate counts of mortality by county for three age groups, they estimated mortality risk associated with air pollution adjusting for age and sex and area-level covariates (education, income level, poverty, and employment), and controlled for potential confounding by cigarette smoking by including standardized mortality ratios for lung cancer and COPD. This study is, therefore, an ecological analysis, similar to past published cross-sectional analyses, in that area-level covariates (education, income level, poverty, and employment) are employed as controlling variables, since individual level information is not available from the Medicare database (other than age and sex), which includes virtually all Americans aged 65 or greater. Exposures are also ecological in nature, as central site data are used as indices of exposure. These results indicated that a 10 µg/m³ increase in the yearly average PM_{2.5} concentration is associated with 10.9% (95% CI: 9.0-12.8) and with 20.8% (95% CI: 14.8-27.1) increases in all-cause mortality for the American Cancer Society and Harvard Six Cities study counties, respectively. The estimates are somewhat higher than those reported by the original investigators, and several possible explanations for this apparent increase are posited by the authors, especially that this is an older population than the ACS cohort. Perhaps the most likely is that the lack of personal confounder information (e.g., past personal smoking information) led to an insufficient control for the effects of these other variables' effects on mortality, inflating the pollution effect estimates somewhat, similar to what has been found in the ACS analyses when only ecological-level control variables were included. The ability of the Eftim *et al.* (2008) study results to qualitatively replicate the original individual-level cohort study (e.g., ACS and Six Cities Study) results suggests that past ecological cross-sectional mortality study results may also provide useful insights into the nature of the association, especially when used for consideration of time trends, or for comparisons of the relative (rather than absolute) sizes of risks between different pollutants or PM components in health effects associations.

The ISA discusses several limitations to this study. The ISA does not, however, discuss that this study may suffer from similar limitations as studies of the SCS and ACS cohorts, and does not actually give additional support for an association between long-term PM_{2.5} exposure and mortality. That is, it is possible that in using similar study methods, the same types of biases and confounders affected results. For example, this study did not adjust for other pollutants, and this likely biased results away from the null. No information was available about the linkage between centrally located air-quality monitors and residents' addresses. Individualized measures of sociodemographic factors (*e.g.*, age, sex, income, smoking) were also not available for the Medicare cohort. In addition, exposure misclassification could have biased results in either direction.

The Medicare cohorts were only followed for three years and PM_{2.5} exposure was measured in that same period. Without prior information on past addresses and exposure levels, a meaningful exposure history that reflects the time course of chronic diseases cannot be constructed. It is inappropriate to attribute present-day deaths to concurrent exposures to PM_{2.5}.

The lack of control for spatially correlated, unmeasured confounders is a major statistical limitation for epidemiologic studies that compare adjusted mortality rates with long-term air pollution exposures across different locations. The ACS re-analysis by HEI confirmed that while the original results held true, the CIs were larger after adjusting for spatial correlation. Eftim *et al.* (2008) acknowledged that their study could also have been affected by this.

Although the ISA states that this type of study is useful for comparison of "relative (rather than absolute)" risks, it also states that findings from this study provide indications that magnitude of association is larger than previously estimated. This is counter-intuitive and suggests that results of this study maybe particularly biased away from the null. Because of these study limitations, this study is not supportive of an association between PM_{2.5} exposures below the current NAAQS and increased mortality.

5.3.6 Zeger *et al.* (2008) based risk estimates on exposures measured with central monitors and Medicare data, did not account for co-pollutants, and found results to vary geographically.

The ISA cites the study by Zeger *et al.* (2008)¹ as showing a statistically significant association between PM_{2.5} exposure and all-cause mortality. Zeger *et al.* (2008) linked Medicare mortality data to

¹ Referenced as Zeger *et al.* (2007) in the ISA.

PM_{2.5} monitoring data, using the same general approach as Eftim *et al.* (2008), to create the Medicare Cohort Air Pollution Study (MCAPS). The study population consisted of 13.2 million Medicare enrollees residing in 4,568 urban zip codes having geographic centroids within six miles of an air monitoring station. The relationship between six-year average exposure to PM_{2.5} and mortality risk in this cohort was assessed over the period from 2000-2005 using log-linear regression models. Median PM_{2.5} concentrations across zip codes were 13.2 µg/m³ (interquartile range: 11.1-14.9). This study indicated that a 10 µg/m³ increase in six-year average PM_{2.5} concentrations was associated with a 6.8% (95% CI: 4.9-8.7) and 13.2% (95% CI: 9.5-16.9) increase in mortality in the eastern and central regions of the US, respectively, when adjusted for SES and COPD. There was no statistically significant association between PM_{2.5} and mortality for zip codes in the western region of the US.

Like the Eftim *et al.* (2008) study, this was an ecological study and prone to exposure misclassification and confounding. A composite area-level SES metric was created, but it does not compensate for potential confounding attributable to the irregular dimensions of the typical zip-code area and within-area heterogeneity in behaviors and SES. As a result of irregular shape, a zip code's centroid may be a point on a map that is far closer to or farther away from an air monitoring station than individuals' residences. There is no way to predict the effect of this potential residual confounding on the calculated risks, and this introduces a significant layer of statistical uncertainty.

Zeger *et al.* (2008) used data from the National Center for Health Statistics (NCHS) to calculate SMRs for COPD for each county. They used these SMRs as surrogates for area-level smoking because direct data on smoking prevalence was unavailable. This created three issues: (1) the potential for ecologic bias from smoking could not be directly evaluated; (2) an inconsistency was created, as the COPD SMRs were at the county level while other covariates were at the zip code-level; and (3) COPD has been associated with PM in prior studies, meaning a potential outcome is being treated as a confounder within the same model.

Importantly, like several other studies reviewed here, the Zeger *et al.* (2008) study did not consider exposure to SO₂ or any other air pollutant that may be correlated with PM_{2.5} and is also associated with mortality. This could have led to results being biased away from the null. Similar to the Eftim *et al.* (2008) study, this Medicare-based cohort had only a six-year window for exposure and outcome measures. Without prior information on past addresses and exposure levels, a meaningful exposure history could not be constructed. As stated above, attributing present-day deaths to present-day exposures to PM_{2.5} is inappropriate.

The statistically significant findings in the eastern and central regions, but not in the western region, are a biological inconsistency. The authors suggest this could be because of higher concentrations of PM and lower mortality in Los Angeles attributable to an unknown and unmeasured protective factor. It is also possible that this is due to higher sulfate levels in PM in the eastern US and the higher nitrate levels in PM in the western US.

Because of the study limitations discussed above, this study is not appropriate for determining whether there is an association between PM_{2.5} exposures below the current annual NAAQS and increased mortality.

5.3.7 Pope *et al.* (2009) found a correlation between PM_{2.5} and life expectancy that could have been explained by other factors, such as NO₂.

The most recent study of the ACS cohort by Pope *et al.* (2009) was published after the first external review draft of the ISA came out, but it is likely it will be included in future drafts. Pope *et al.* (2009) analyzed the association between life expectancy and average ambient PM_{2.5} levels in 211 county units in 51 US metropolitan areas between the time periods 1979-1983 *versus* 1999-2000. To obtain exposure data for the earlier time interval, the authors used information from the re-analysis and extended analysis of the ACS prospective cohort study. For the latter time interval, data were extracted from the EPA's AIRS database. For sensitivity analyses, the authors used various surrogate or proxy measures: per capita income and high-school graduation status for health indicators and death rates from lung cancer and COPD for smoking prevalence. As would be expected, life expectancy increased between 1980 and 2000 (2.7 years), and ambient PM_{2.5} decreased (6.5 µg/m³). The authors correlated changes in life expectancy to matching data in PM_{2.5} reductions. After adjustment for various socioeconomic and demographic factors, the authors concluded that, out of the overall increase in life expectancy, the slope of this correlation was such that a decrease of 10 µg/m³ in PM_{2.5} is associated with an increase in life expectancy of 0.61 year. They state that PM_{2.5} decreases "accounted for as much as 15% of the overall increase in life expectancy."

This is an ecological study, with no information on personal exposure or individual outcome, and inferences about the nature of specific individuals are based on aggregate statistics collected for the geographic area in which the individuals live. The graph of the change in life expectancy *versus* the change in PM_{2.5} shows extensive scatter around the regression line for each locale. Even after adjusting

for socioeconomic and demographic variables and proxy variables for smoking, the variations in changes of life expectancy were as great as 10-fold. The authors acknowledged, but did not discuss, that other factors in addition to air pollution were influencing changes in life expectancy. During the two decades that this study covered, improvements in overall health care, early detection programs, more effective pharmaceuticals, health education and awareness, chronic disease management, dietary awareness, and overall standard of living all contributed to increased life expectancy. To the extent that the contemporaneous co-variation of these changes cannot be untangled from the changes in PM_{2.5}, they remain as potentially strong confounders.

The authors note that the three variables (per capita income for life style factors, lung cancer and COPD death rates for smoking prevalence, and PM_{2.5} for co-pollutants) that were most strongly associated with changes in life expectancy were all proxy measures. Thus, more attention should be paid to what each variable, in fact, represents. For example, PM_{2.5} is only one of the criteria pollutants, and these other pollutants, as well as the 180 hazardous air pollutants (HAPs) have likely also decreased during the 20 years covered by the this study. Thus, it is highly unlikely that the correlation between increased life expectancy and decreased PM_{2.5} is attributable to PM_{2.5} alone. The authors acknowledged this point, but disregarded the quantitative importance of specifically designating PM_{2.5} alone in their analyses.

Because of these limitations, this study is not appropriate for determining whether there is an association between PM_{2.5} exposures below the current annual NAAQS and life expectancy.

5.4 Long-term studies of reproductive and developmental effects do not provide support for a causal association.

Section 7.4 of the ISA evaluates the scientific evidence for associations between PM and reproductive or developmental outcomes. The ISA states that the epidemiological evidence is suggestive of a causal relationship between long-term exposures to PM_{2.5} and reproductive and developmental outcomes. The ISA also states that epidemiologic studies do not consistently report associations between PM exposure and reproductive or developmental outcomes, but studies with positive associations are often emphasized over those that report no association. For example, in Section 7.4.1, the ISA compared a study by Maisonet *et al.* (2001), which showed no increased risk for low birth weight associated with PM₁₀ exposure, to a study by Bell *et al.* (2007), which showed that reductions in birth weight were

associated with exposure to both PM_{2.5} and PM₁₀. The ISA noted that a "larger sample size was able to detect a small increase in risk" in the Bell *et al.* (2007) study, but this is an insignificant point because both studies relied on ambient monitoring as a surrogate for personal exposure and, thus, exposure misclassification likely occurred in both studies. The ISA also stated that misclassification is reduced in particular studies that only include women living within five miles of a monitoring station, and that such a reduction in misclassification should lead to a stronger association. Such a hypothesis is unverified, however, and it is unclear how reliance on an exposure metric up to five miles away could significantly reduce exposure misclassification. In a discussion of the inconsistent results among international studies, the ISA further discounts studies that do not find a positive association by stating: "Studies with negative results must be interpreted with caution when comparison groups have significant exposure." Many of the studies upon which the ISA relied for a causal determination for reproductive and developmental outcomes suffer from limitations that limit their usefulness as evidence for an association between these outcomes and PM at levels below the current NAAQS. Two of these studies are described below.

5.4.1 Bell *et al.* (2007) based risk estimates on exposures measured with central monitors and did not rule out NO₂ or several other confounders.

Bell *et al.* (2007) investigated the effects of PM and other air pollutants on birth weight in Connecticut and Massachusetts over a four-year period (1999 through 2002). Exposures were assigned to each pollutant over the gestational period and each trimester using the county of mother's residence at delivery. Covariates such as marital status, tobacco use during pregnancy, education level, and other previously identified risk factors for low birth weight were included. Average PM_{2.5} concentrations in this study were below the NAAQS. The results indicated that a 2.2 µg/m³ increase in PM_{2.5} was associated with a difference in birth weight of -14.7 g (95% CI: -17.1 to -12.3) in a linear model, and with an odds ratio of 1.054 (95% CI: 1.022 to 1.087) in a logistic model. The associations were robust to inclusion of other pollutants whose exposures were not correlated with PM_{2.5}. Exposures to PM_{2.5} and NO₂ were highly correlated in the data set, however, so this study could not distinguish between the effects of these pollutants.

Other limitations of the Bell *et al.* (2007) study include the use of county-wide ambient monitoring as a surrogate for personal exposure, as discussed above, which can lead to exposure misclassification. This is especially true if the county of the mother's residence is not constant throughout gestation. The adjustments made for education level may not fully address potential confounding by SES because numerous other factors are involved. Tobacco use was recorded as a binary measure (yes/no),

which may not have fully captured its effects; number of cigarettes smoked during pregnancy or time of occurrence during the pregnancy (particularly in a vulnerable period) would be more informative and could help to decrease the bias of associations that can result from misclassification of a covariate. Although the associations between PM_{2.5} and low birth weight observed in this study were statistically significant and occurred at levels below the current NAAQS, the limitations of the study, particularly with respect to confounding by co-pollutants and other variables, are such that the results should be interpreted with caution.

5.4.2 Parker *et al.* (2008) did not measure PM_{2.5} or exclude preterm births as a result of medical intervention from their study.

Parker *et al.* (2008) examined preterm birth in mothers in the Utah Valley, where an open-hearth steel mill was closed for a 13-month period in 1986-1987. Air pollution, particularly PM₁₀, was known to be reduced during the closure. Birth outcomes for mothers residing within and outside of the Utah Valley were compared before, during, and after the mill closure. The authors reported a relative risk of 0.86 (95% CI: 0.75 to 0.98) for preterm birth in women residing in the Utah Valley who were in the second trimester of pregnancy during the mill closure. Decreased risks were also observed during the mill closure for other periods of pregnancy, but these were not statistically significant. Risks did not change for women who resided outside of the Utah Valley during the same time period.

Parker *et al.* (2008) acknowledged that their study was limited by the small number of births and imprecise exposure assessments. Indeed, no direct measure of PM was obtained, and exposures to "air pollution" were simply categorized as reduced (during the period of mill closure) or not. Thus, the pollutants with the strongest impact on risk could not be identified in this study. In addition, potential misclassification of preterm birth is a concern in this study, because the authors did not distinguish between preterm birth that occurs naturally or as a result of medical intervention, and exposure may be associated with only one of these types of preterm birth. This potential for misclassification and the lack of information regarding which air pollutant(s) may have impacted risk in this study make it unsuitable for use as evidence that exposure to PM is associated with an increased risk of preterm birth.

5.5 Selected studies investigating the effects of long-term exposure to PM_{2.5} do not support a causal association with morbidity or mortality at levels below the current NAAQS

In the ISA, the conclusion by US EPA that there is a likely causal relationship between long-term exposure to PM_{2.5} and cardiovascular morbidity appears to be based on research of short-term exposure. The ISA states that the results were inconsistent in long-term exposure studies, but "the evidence from epidemiologic, human clinical, and animal toxicological studies that examined the cardiovascular outcomes associated with short-term exposure to PM_{2.5} (discussed in Section 6.2), supports a role for the development of cardiovascular morbidity in response to long-term exposure to PM_{2.5}. Based on the consistent and coherent evidence from epidemiologic and toxicological studies that examined the association between long-term and short-term exposure to PM_{2.5} and cardiovascular morbidity, sufficient evidence is available to conclude that a causal relationship is likely to exist between long-term exposure to ambient concentrations of PM_{2.5} and cardiovascular morbidity." The ISA states that the association between PM_{2.5} and respiratory morbidity is likely to be causal because "collectively," toxicological studies provide biological plausibility and "overall," evidence from epidemiological and toxicological studies is consistent and coherent. The ISA also states: "The new epidemiologic evidence reports a consistent association between long-term exposure to PM_{2.5} and an increased risk of mortality (with the majority of the effects ranging from > 1 to 1.20) in cities with annual average PM_{2.5} concentrations ranging from 10.2-29 µg/m³ (see Section 7.6)."

I reviewed four studies on which US EPA based its conclusions regarding long-term PM_{2.5} exposure and cardiovascular morbidity: two of which analyzed subclinical effects (Allen *et al.*, 2009; Diez Roux *et al.*, 2008), and two of which analyzed clinical outcomes in epidemiological studies (Hoffman *et al.*, 2006; Miller *et al.*, 2007). I also reviewed one study, by Goss *et al.* (2004), that US EPA relied on for its assessment of respiratory morbidity, seven of the major studies published since the 2004 AQCD that examined PM_{2.5} exposure and mortality (Beleen *et al.*, 2008; Eftim *et al.*, 2008; Jerrett *et al.*, 2005; Laden *et al.*, 2006; Miller *et al.*, 2007; Pope *et al.*, 2009; Zeger *et al.*, 2008), and two studies relied on in the ISA for assessment of reproductive and developmental outcomes (Bell *et al.*, 2007; Parker *et al.*, 2008).

Based on my review of these long-term exposure studies, I conclude that the new epidemiology studies do not support a causal association between PM_{2.5} at concentrations below the current NAAQS

and health effects. That is, the long-term exposure studies of PM_{2.5} should not be used as evidence of a "likely" causal relationship between ambient PM_{2.5} at concentrations below the current NAAQS and cardiovascular morbidity, respiratory morbidity, or mortality, nor as "suggestive" evidence of a causal relationship between PM_{2.5} below the current NAAQS and reproductive and developmental outcomes.

Several of these studies reported either null or weakly positive findings (Allen *et al.*, 2009; Diez Roux *et al.*, 2008). In other cases, weakly positive findings became non-significant when adjusted for confounders (*e.g.*, Allen *et al.*, 2009). Several studies did not have information on co-pollutants or other factors that may have been associated with exposure and/or outcome (such as people living closer to monitors may have low SES and be at higher risks for certain outcomes), so reported associations were likely biased away from the null (Allen *et al.*, 2009; Diez Roux *et al.*, 2008; Beelen *et al.*, 2008; Eftim *et al.*, 2008; Jerrett *et al.*, 2005; Laden *et al.*, 2006; Miller *et al.*, 2007; Pope *et al.*, 2009; Zeger *et al.*, 2008).

Exposure misclassification/measurement error is perhaps the biggest shortcoming of long-term studies of PM_{2.5}. All studies reviewed here used measurements from central monitors and, because the distance between people's residences to these monitors varied, this led to inaccurate measurements. In addition, some studies used exposure measurements from 2000 to represent earlier exposures (*e.g.*, Jerrett *et al.*, 2005). As PM_{2.5} concentrations have been decreasing over time, this likely overestimated risks, particularly when studies examine risks with small increments of exposure (*e.g.*, 10 µg/m³). Other studies estimated exposure to PM_{2.5} based on measurements of PM₁₀ exposures, and this could bias results in either direction (Beelen *et al.*, 2008; Laden *et al.*, 2000).

All of these factors make it inappropriate to attribute risks to long-term exposure to PM_{2.5}. More importantly, the long-term studies relied on in the ISA had few exposures below 15 µg/m³, so they were not informative regarding risks below the current NAAQS.

6 Chapter 8 does not accurately portray the health impacts of PM_{2.5} exposure.

In Chapter 8, the PM ISA addresses the concentration-response relationship for PM and key health effects, as well as PM concentrations at which health effects occur, including in sensitive and vulnerable individuals. The ISA presents studies and analyses that suggest the concentration-response for PM and key health effects, in particular morbidity and mortality, may be linear, without a threshold. The ISA does not, however, fully consider evidence that is not consistent with a linear no-threshold concentration-response model for PM and morbidity or mortality.

6.1 Heterogeneity among cities/regions may influence the concentration-response relationship.

In the ISA, US EPA relied upon multi-city studies with combined concentration-response curves to evaluate the relationship between PM exposure and mortality. These studies were conducted by Daniels *et al.* (in HEI, 2004), Schwartz (2004), and Samoli *et al.* (2005) for short-term exposures and by Schwartz *et al.* (2008) and Roman *et al.* (2008) for long-term exposures, and are discussed in more detail below, in Section 6.4.

The ISA acknowledged that "the heterogeneity observed between cities complicates the biological explanation for the combined and city-specific results." Indeed, individual cities and regions differ in ways that might result in different concentration-response relations, including differences in the sources, composition, and concentrations of PM; differences in the adequacy of monitoring networks in measuring population exposure; and differences in population characteristics that may lead to differences in population susceptibility to PM effects (HEI, 2004). All of these differences would likely exist among the cities or regions and may affect estimates of concentration-response relationships across cities. Such heterogeneity may influence the shape, as well as the slope, of the relationship, and could make the concentration-response curve appear linear at low exposure levels when in reality it is not. Because of these issues, consistency of findings across cities is important in evaluating causality in the concentration-response relationship. Tests used for heterogeneity tend to have low power, however, so even if heterogeneity is not statistically significant in certain studies, it might still be present (HEI, 2004). Thus, it is difficult to give meaning to combined-city concentration-response curves described in several studies

(e.g., Laden *et al.*, 2006; Miller *et al.*, 2007) in the face of possible between-city or between-region heterogeneity in the shape of the curve.

6.2 Heterogeneity among the population will not linearize the concentration-response relationship.

The ISA states that at the human population level, individual differences in susceptibility to air pollution health effects tend to smooth and "linearize" the concentration-response function. Thus, inter-individual variability can complicate the ability to determine the shape of the PM concentration-response curve and the potential presence of a threshold. This inter-individual variability could be due to the fact that each individual has a different threshold. It has been suggested that the variation in these individual thresholds necessarily leads to a non-threshold concentration-response model for the population (White *et al.*, 2008). This is not the case, however. Heterogeneity in sensitivity and in modifying factors among people in the target population may broaden the concentration-response relationship, but does not linearize it. The combined effect of variation in many modifying factors leads to the expectation of a cumulative lognormal dose-response function (which is always nonlinear) rather than a linear one (Rhombert, 2009). Heterogeneity among the populations studied may have contributed to a misinterpretation of the studies relied upon by US EPA to evaluate the concentration-response relationship between mortality and PM exposure.

6.3 Measurement error can artificially flatten concentration-response curves.

Estimating individual exposure to air pollutants from central-site outdoor pollution monitors may result in considerable error (Brauer *et al.*, 2002). Some individuals in the population will have greater exposures than others for any given central-site ambient concentration, which will broaden the normal distribution of risks due to inter-individual variability, as discussed above. Even in the unlikely event in which all individuals in a population have the same true concentration-response threshold, misclassification of exposure could result in some individuals appearing to be affected below this threshold and others appearing not to be affected even above their true threshold. Thus, thresholds in the relationship between individual exposure to PM and risks of morbidity and mortality may be blurred by measurement error when studied by using the relationships between concentrations at central-site monitors and aggregate morbidity or mortality (HEI, 2004). In fact, exposure measurement error may artificially flatten apparent concentration-response curves and tend to make any concentration-related

effect (even those that are truly threshold in nature) look more or less linear as an artifact of the analysis, thus masking what may in fact be a threshold.

Several studies have addressed the effect of measurement error on the concentration-response relationship. Schwartz and Zanobetti (2000) argued that it is possible to detect threshold relations in meta-analyses down to low concentration levels, but they did not directly address the effect of measurement error on concentration-response shapes within individual cities upon which the meta-analyses were based. Cakmak *et al.* (1999) dealt only with population-level data and did not directly address the impact of individual measurement error on the shape of the concentration-response curve. Brauer *et al.* (2002) used measured ambient and personal PM concentrations to demonstrate that measurement error can either underestimate the threshold concentration or obscure the threshold altogether. This is because measurement error can bias the magnitude of effect estimates, lead to biased regression coefficients, and affect the ability to observe a threshold level, should one exist (Brauer *et al.*, 2002). These investigators showed that if exposure misclassification is reduced by the use of appropriate exposure metrics, then common underlying individual thresholds result in similar population level thresholds. Their simulations assumed that all individuals in a population have the same threshold concentration and the same slope of their concentration-response curves, but it was shown that the obscuring of thresholds would be even greater if the simulations incorporated thresholds that varied across individuals. Thus, the simulations of Brauer *et al.* (2002) suggest that the inability to detect a threshold in many epidemiological studies does not, in fact, mean that no threshold exists.

The possibility that exposure measurement error obscures thresholds limits the ability to draw conclusions about the absence of a threshold in the PM studies relied on in the ISA. These studies are discussed below, in Section 6.4.

6.4 Studies relied on in the ISA to assess concentration-response relationships are not sufficient for concluding a linear model.

The ISA points out that a multitude of factors have been identified that complicate the ability to determine the shape of the PM concentration-response curve and the potential presence of a threshold, including inter-individual variability; additivity of pollutant-induced effects to naturally occurring background disease processes; exposure error; response error; and low data density in the lower concentration range. Despite these factors, the ISA suggests that linear no-threshold models best describe

the associations between PM and health effects. Evidence suggests that these may not be the best models to describe this association, however.

For evaluating the concentration-response relationship between mortality and short-term exposure to PM, US EPA relied on studies by Daniels *et al.* (in HEI, 2004), Schwartz (2004), and Samoli *et al.* (2005).

Daniels *et al.* (as cited in HEI, 2004) analyzed three possible models to describe the relationship between PM₁₀ and mortality and concluded that a log-linear model was the most appropriate model for both cardiorespiratory and total mortality (but not other-cause mortality). The models were analyzed using Akaike Information Criterion (AIC), however, and this may not be an appropriate criterion upon which to base the choice between models. The AIC is a way of choosing among models for the one that is most explanatory of the variation in the dependent variable among study subjects, with a "penalty" for models that explain more just by adding parameters. It is aimed at helping to decide when added model flexibility, by adding parameters, is really explaining the underlying true causal factors better, and not just over-fitting errors. It can be used to be parsimonious – to argue whether a more complex model shape better describes the data in the model (which may or may not accurately portray the "truth") – but it is not really an appropriate criterion for whether there actually is or is not more model curvature in the "true" model. Thus, it is inappropriate to use the AIC to determine whether the linear no-threshold model best describes the relationship between PM and health effects. Although the ISA acknowledges the limitations of using the AIC, it bears mentioning here because the Daniels *et al.* (2004) study is relied on by US EPA and this discussion may have implications for other studies.

The HEI (2004) notes that another issue relating to the Daniels *et al.* (in HEI, 2004) study is their determination that the relationship between PM₁₀ and both cardiovascular-respiratory mortality and total mortality is log linear, but the relationship between PM₁₀ and other-cause (besides cardiovascular-respiratory) mortality is not log linear. The concentration-response curve for total deaths must logically be the average of the curves for cardiovascular-respiratory and other-cause mortality, weighted by the proportion of deaths in each. If the association of PM₁₀ with "other" deaths is nonlinear, the association with total deaths is unlikely to be log linear and, indeed, cannot be log linear if the association with cardiovascular-respiratory deaths is log linear. Thus, the lack of evidence against linearity from analyses of total deaths may reflect loss of power in analyses of this less specific outcome (HEI, 2004).

Schwartz (2004) analyzed the relationship between short-term exposure to PM and mortality using a different technique than Daniels *et al.* (in HEI, 2004), including indicator variables for days in which the PM₁₀ concentration was within specific ranges. Schwartz (2004) did not find evidence for non-linearity when combining risk estimates across 14 cities. This study did not analyze city-specific thresholds, however, and heterogeneity in the concentration-response curve across cities was not examined. As noted above, heterogeneity across cities may influence the shape and slope of the concentration-response relationship, making it appear linear at low exposure levels when it is actually non-linear.

Samoli *et al.* (2005) observed heterogeneity in the shape of the concentration-response curve across 22 European cities. Their analysis supported a log-linear association between PM₁₀ and mortality, but the ISA correctly stated that "the heterogeneity observed between cities complicates the biological explanation for the combined and city-specific results."

Regarding the aforementioned studies, the ISA states that they "all support the use of a no-threshold log-linear model, but additional issues such as the influence of heterogeneity in estimates between cities, and the effect of seasonal and regional differences in PM on the C-R relationship still require further investigation."

For mortality associated with long-term exposure to PM, the ISA relied on studies by Schwartz *et al.* (2008) and Roman *et al.* (2008). Schwartz *et al.* (2008) examined the concentration-response relationship between PM_{2.5} and mortality using data from the Harvard Six Cities Study. Two approaches were used, each involving Cox proportional hazards models, and both approaches found that the concentration-response curve was indistinguishable from linear. Yet, the use of the Cox model may have led to biased results (see Section 2.1.4).

Roman *et al.* (2008) developed probabilistic uncertainty distributions to characterize uncertainties in the concentration-response relationship for annual PM concentrations ranging from 4 to 30 µg/m³. A panel of 12 experts was asked to provide judgment on the true shape of the concentration-response curve. The majority of the panel agreed that, collectively, the epidemiologic data did not provide evidence of a population threshold. Several of these experts were authors of key air pollution studies, however, so they may have had preconceived opinions regarding the nature of the concentration-response relationship. Other underlying cognitive tendencies that influence expert judgment, but cannot be accounted for, include the tendency to assign greater probability to frequently mentioned events, the tendency to be

over-influenced by the first pieces of information provided, and the tendency of experts to overestimate the probability that their opinions are correct. In addition, the study by Roman *et al.* (2008) emphasized the conclusions of the expert panel, but not the data that went into these conclusions. Thus, one cannot evaluate the appropriateness of the analyses.

In the epidemiology studies relied on by the ISA, many of the uncertainties within the data (such as confounding, measurement error, and exposure misclassification) were not accounted for in the statistical models. Even if a linear model best describes the reported data, it is plausible that a non-linear model would have better described the data were these uncertainties taken into account. Thus, the currently-available PM epidemiology data are simply not robust enough to determine whether a linear no-threshold model best describes the association between PM exposure and health effects.

7 Conclusions

In its last review of PM_{2.5}, US EPA (2004) concluded that exposure to ambient PM caused or was associated with a wide variety of health effects. Further, US EPA concluded that no threshold had been identified below which these health effects occur. The ISA does not adequately demonstrate that the studies published since the last review: (1) demonstrate that PM_{2.5} causes additional health effects not identified in the last review; (2) provide reduced uncertainties or stronger evidence for the previously identified effects; (3) provide evidence that risk estimates for the previously identified effects have increased since the last review; or (4) provide further information on the possibility that these effects occur at lower levels than previously identified.

References

- Abrahamowicz, M; Schopflocher, T; Leffondre, K; du Berger, R; Krewski, D. 2003. "Flexible modeling of exposure-response relationship between long-term average levels of particulate air pollution and mortality in the American Cancer Society study." *J. Toxicol. Environ. Health* 66(16-19):1625-1654.
- Allen, RW; Criqui, MH; Diez Roux, AV; Allison, M; Shea, S; Detrano, R; Sheppard, L; Wong, ND; Stukovsky, KH; Kaufman, JD. 2009. "Fine particulate matter air pollution, proximity to traffic, and aortic atherosclerosis." *Epidemiology* 20(2):254-264.
- Beelen, R; Hoek, G; van den Brandt, PA; Goldbohm, RA; Fischer, P; Schouten, LJ; Jerrett, M; Hughes, E; Armstrong, B; Brunekreef, B. 2008. "Long-term effects of traffic-related air pollution on mortality in a Dutch cohort (NLCS-AIR Study)." *Environ. Health Perspect.* 116:196-202.
- Bell, ML; Ebisu, K; Belanger, K. 2007. "Ambient air pollution and low birth weight in Connecticut and Massachusetts." *Environ. Health Perspect.* 115:1118-1125.
- Bell, ML; Ebisu, K; Peng, RD; Walker, J; Samet, JM; Zeger, SL; Dominici, F. 2008. "Seasonal and regional short-term effects of fine particles on hospital admissions in 202 US counties, 1999-2005." *Am. J. Epidemiol.* 168(11):1301-1310.
- Boffetta, P; McLaughlin, JK; La Vecchia, C; Tarone, RE; Lipworth, L; Blot, WJ. 2008. "False-positive results in cancer epidemiology: A plea for epistemological modesty." *J. Natl. Cancer Inst.* 100:988-995.
- Brauer, M; Brumm, J; Vedal, S; Petkau, AJ. 2002. "Exposure misclassification and threshold concentrations in time series analyses of air pollution health effects." *Risk Anal.* 22(6):1183-1193.
- Brown, JS; Zeman, KL; Bennett, WD. 2001. "Regional Deposition of Coarse Particles and Ventilation Distribution in Healthy Subjects and Patients with Cystic Fibrosis." *J Aerosol Med* 14:443-454.
- Bukowski, J. 2008. "Do Pollution Time-Series Studies Contain Uncontrolled or Residual Confounding by Risk Factors for Acute Health Events?" *Reg. Tox. & Pharm.* 51:135-140.
- Bukowski, J. 2007. "Do Time-Series Studies Contain Residual Confounding by Risk Factors for Acute Health Events?" *Epidemiology* 18:839.
- Bukowski, JA. 2008. "Response to counterpoint: Time-series studies of acute health events and environmental conditions are not confounded by personal risk factors by Mark S. Goldberg, Richard T. Burnett, Jeffrey R. Brook." *Regul. Toxicol. Pharmacol.* 51:148-150.
- Burnett, RT; Stieb, D; Brook, JR; Cakmak, S; Dales, R; Raizenne, M; Vincent, R; Dann, T. 2004. "Associations between short-term changes in nitrogen dioxide and mortality in Canadian cities." *Arch. Environ. Health* 59(5):228-236.
- Cakmak, S; Burnett, RT; Krewski, D. 1999. "Methods for detecting and estimating population threshold concentrations for air pollution-related mortality with exposure measurement error." *Risk Anal.* 19(3):487-496.

- Demchuk, E; Yucesoy, B; Johnson, VJ; Andrew, M; Weston, A; Germolec, DR; De Rosa, CT; Luster, MI. 2007. "A statistical model for assessing genetic susceptibility as a risk factor in multifactorial diseases: Lessons from occupational asthma." *Environ. Health Perspect.* 115(2):231-234.
- Diez Roux, AV; Auchincloss, AH; Franklin, TG; Raghunathan, T; Barr, RG; Kaufman, J; Astor, B; Keeler, J. 2008. "Long-term exposure to ambient particulate matter and prevalence of subclinical atherosclerosis in the Multi-Ethnic Study of Atherosclerosis." *Am. J. Epidemiol.* 167(6):667-675.
- Dominici, F; Peng, RD; Bell, ML; Pham, L; McDermott, A; Zeger, SL; Samet, JM. 2006. "Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases." *JAMA* 295(10):1127-1134.
- Dominici, F; Peng, RD; Zeger, SL; White, RH; Samet, JM. 2007. "Particulate air pollution and mortality in the United States: Did the risks change from 1987 to 2000?" *Am. J. Epidemiol.* 166(8):880-888.
- Eftim, SE; Samet, JM; Janes, H; McDermott, A; Dominici, F. 2008. "Fine particulate matter and mortality: A comparison of the Six Cities and American Cancer Society cohorts with a Medicare cohort." *Epidemiology* 19:209-216.
- Franklin, M; Zeka, A; Schwartz, J. 2007. "Association between PM-2.5 and all-cause and specific-cause mortality in 27 US communities." *J. Expo. Sci. Environ. Epidemiol.* 17:279-287.
- Glymour, MM; Greenland, S. 2008. "Causal Diagrams." In *Modern Epidemiology (Third Edition)*. (Eds.: Rothman, KJ; Greenland, S; Lash, TL), Lippincott, Williams, & Wilkins, Philadelphia, PA.
- Goldberg, MS; Burnett, RT; Brook, JR. 2008. "Counterpoint: Time-series studies of acute health events and environmental conditions are not confounded by personal risk factors." *Regul. Toxicol. Pharmacol.* 51:141-147.
- Gong, H; Linn, WS; Sioutas, C; Terrell, SL; Clark, KW; Anderson, KR; Terrell, LL. 2003. "Controlled exposures of healthy and asthmatic volunteers to concentrated ambient fine particles in Los Angeles." *Inhal. Toxicol.* 15(4):305-325.
- Gong, H; Linn, WS; Terrell, SL; Anderson, KR; Clark, KW; Sioutas, C; Cascio, WE; Alexis, N; Devlin, RB. 2004. "Exposures of elderly volunteers with and without chronic obstructive pulmonary disease (COPD) to concentrated ambient fine particulate pollution." *Inhal. Toxicol.* 16(11-12):731-744.
- Goss, CH; Newsom, SA; Schildcrout, JS; Sheppard, L; Kaufman, JD. 2004. "Effect of ambient air pollution on pulmonary exacerbations and lung function in cystic fibrosis." *Am. J. Respir. Crit. Care Med.* 169:816-821.
- Health Effects Institute (HEI). 2004. "The National Morbidity, Mortality, and Air Pollution Study. Part III: PM10 Concentration-Response Curves and Thresholds for the 20 Largest US Cities." HEI Research Report 94 Part III. 30p., May.
- Hill, AB. 1965. "The environment and disease: Association or causation?" *Proc. R. Soc. Med.* 58:295-300.

Hoffmann, B; Moebus, S; Stang, A; Beck, EM; Dragano, N; Mohlenkamp, S; Schmermund, A; Memmesheimer, M; Mann, K; Erbel, R; Jockel, KH; Heinz Nixdorf RECALL Study Investigative Group. 2006. "Residence close to high traffic and prevalence of coronary heart disease." *Eur. Heart J.* 2:2696-2702.

Institute of Medicine. 2008. *Improving the Presumptive Disability Decision-Making Process for Veterans*. (Eds.: J.M. Samet; C.C. Bodurow), The National Academies Press, Washington, DC. Accessed at http://books.nap.edu/openbook.php?record_id=11908&.

Jerrett, M; Burnett, RT. 2007. "Air pollution and cardiovascular events (Letter)." *N. Engl. J. Med.* 356(20):2105.

Jerrett, M; Burnett, RT; Ma, R; Pope, CA, 3rd; Krewski, D; Newbold, KB; Thurston, G; Shi, Y; Finkelstein, N; Calle, EE; Thun, MJ. 2005. "Spatial analysis of air pollution and mortality in Los Angeles." *Epidemiology* 16:727-736.

Jurek, AM; Greenland, S; Maldonado, G. 2008. "How far from non-differential does exposure or disease misclassification have to be to bias measures of association away from the null?" *Int. J. Epidemiol.* 37(2):382-385.

Jurek, AM; Greenland, S; Maldonado, G; Church, TR. 2005. "Proper interpretation of non-differential misclassification effects: Expectations vs. observations." *Int. J. Epidemiol.* 34(3):680-687.

Laden, F; Schwartz, J; Speizer, FE; Dockery, DW. 2006. "Reduction in fine particulate air pollution and mortality: Extended follow-up to the Harvard Six Cities Study." *Am. J. Respir. Crit. Care Med.* 173(6):667-672.

Maisonet M; Bush TJ; Correa A; Jaakkola JJ. (2001). Relation between ambient air pollution and low birth weight in the Northeastern United States. *Environ Health Perspect* 109(Suppl 3):351-356.

Miller, KA; Siscovick, DS; Sheppard, L; Shepherd, K; Sullivan, JH; Anderson, GL; Kaufman, JD. 2007. "Long-term exposure to air pollution and incidence of cardiovascular events in women." *N. Engl. J. Med.* 356:447-457.

Moolgavkar, SH. 2005. "A review and critique of the EPA's rationale for a fine particle standard." *Regul. Toxicol. Pharmacol.* 42:123-144.

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

Parker, JD; Mendola, P; Woodruff, TJ. 2008. "Preterm birth after the Utah Valley Steel Mill closure: A natural experiment." *Epidemiology* 19(6):820-823.

Pope, CA; Burnett, RT; Thun, MJ; Calle, EE; Krewski, D; Ito, K; Thurston, GD. 2002. "Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution." *JAMA* 287(9):1132-1141.

Pope, CA; Ezzati, M; Dockery, DW. 2009. "Fine-particulate air pollution and life expectancy in the United States." *N. Engl. J. Med.* 360(4):376-386.

Rhomberg, LR. 2009. "Linear low-dose extrapolation for non-cancer responses is not generally appropriate." *Environ. Health Perspect.* In press.

Roman, HA; Walker, KD; Walsh, TL; Conner, L; Richmond, HM; Hubbell, BJ; Kinney, PL. 2008. "Expert judgment assessment of the mortality impact of changes in ambient fine particulate matter in the U.S." *Environ. Sci. Technol.* 42(7):2268-2274. Accessed on January 30, 2009 at <http://pubs.acs.org/doi/pdf/10.1021/es0713882>.

Samoli, E; Analitis, A; Touloumi, G; Schwartz, J; Anderson, HR; Sunyer, J; Bisanti, L; Zmirou, D; Vonk, JM; Pekkanen, J; Goodman, P; Paldy, A; Schindler, C; Katsouyanni, K. 2005. "Estimating the exposure-response relationships between particulate matter and mortality within the APHEA multicity project." *Environ. Health Perspect.* 113(1):88-95.

Schwartz, J. 2004. "The effects of particulate air pollution on daily deaths: A multi-city case crossover analysis." *Occup. Environ. Med.* 61(12):956-961.

Schwartz, J; Coull, B; Laden, F; Ryan, L. 2008. "The effect of dose and timing of dose on the association between airborne particles and survival." *Environ. Health Perspect.* 116(1):64-69.

Schwartz, J; Zanobetti, A. 2000. "Using meta-smoothing to estimate dose-response trends across multiple studies, with application to air pollution and daily death." *Epidemiology* 11(6):666-672.

Sorahan, T; Gilthorpe, MS. 1994. "Non-differential misclassification of exposure always leads to an underestimate of risk: An incorrect conclusion." *Occup. Environ. Med.* 51(12):839-840.

US EPA, National Center for Environmental Assessment (NCEA) - RTP Division. 2008. "Integrated Science Assessment for Particulate Matter (First External Review Draft)." EPA/600/R-08/139. 938p., December.

US EPA, Office of Research and Development. 2004. "Air Quality Criteria for Particulate Matter. Volume I of II." US EPA, National Center for Environmental Assessment (NCEA) (Research Triangle Park, NC) EPA/600/P-99/002aF. 900p., October.

Valberg, PA. 2003. "Possible noncausal bases for correlations between low concentrations of ambient particulate matter and daily mortality." *Nonlinearity Biol. Toxicol. Med.* 1:521-530.

Wacholder, S; Hartge, P; Lubin, JH; Dosemeci, M. 1995. "Non-differential misclassification and bias towards the null: A clarification [Letter to the editor]." *Occup. Environ. Med.* 52(8):557.

White, RH; Cote, I; Zeise, L; Fox, M; Dominici, F; Burke, TA; White, PD; Hattis, DB; Samet, JM. 2008. "State-of-the-Science Workshop Report: Issues and Approaches in Low Dose-Response Extrapolation for Environmental Health Risk Assessment." Accessed at <http://www.ehponline.org/members/2008/11502/11502.pdf>.

Zeger, SL; Dominici, F; McDermott, A; Samet, JM. 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.