EPA's Assessment of Health Benefits Associated with $PM_{2.5}$ Reductions for the Final Mercury and Air Toxics Standards

Prepared for

The American Energy Initiative Hearing

Congress of the United States
House of Representatives
Committee on Energy and Commerce
Subcommittee on Energy and Power
2125 Rayburn House Office Building
Washington, DC 20515-6115

Prepared by
Julie E. Goodman, Ph.D., DABT

February 8, 2012

Gradient
www.gradientcorp.com
Science and Strategies for Health and the Environment

20 University Road
Cambridge, MA 02138
617-395-5000

Overview

EPA estimated that the Mercury and Air Toxics Standards will reduce the disease burden in America to such an extent that it will translate to tens of billions of dollars saved. The largest benefits from the Mercury and Air Toxics Standards are derived not from reducing mercury, but from reducing fine particulate matter ($PM_{2.5}$). Despite the vast array of peer-reviewed scientific literature on the topic, EPA based its calculations on only two $PM_{2.5}$ epidemiology studies that reported statistical associations between $PM_{2.5}$ reductions and health benefits and assumed a causal relationship. These studies had methodological limitations and were not consistent with many epidemiology studies indicating no correlation between reducing $PM_{2.5}$ and health benefits or experimental studies indicating an exposure threshold below which $PM_{2.5}$ is not likely to overwhelm the body's natural defenses. Thus, EPA's analysis led to grossly inflated estimates of benefits.

My biographical summary is included at the end of this testimony, followed by an Appendix that further details the uncertainties associated with estimations of health benefits from $PM_{2.5}$ reductions.

Testimony

Good morning, Mr. Chairman and members of the subcommittee, and thank you for the opportunity to testify. I am Dr. Julie Goodman, a board-certified toxicologist and Principal at Gradient, an environmental consulting firm in Cambridge, Massachusetts. I also teach a graduate-level epidemiology course at the Harvard School of Public Health. I am presenting testimony this morning on my own behalf as an independent scientist.

I want to start by stressing how important clean air is. There is no doubt that high levels of pollution can be detrimental to human health and the environment. Considering everything from infant mortality to life expectancy, negative impacts from air pollution are at their lowest levels in recent history in the United States.

EPA has estimated that the Mercury and Air Toxics Standards, also known as the Utility MACT, will lead to benefits from reductions in health impacts ranging from bronchitis to mortality, and that these benefits translate to tens of billions of American dollars saved. But the methods used to derive these benefits are fraught with large uncertainties, which likely resulted in a large overestimation of benefits.

Despite its name, the vast majority of the benefits from the Mercury and Air Toxics Standards reported by EPA are not from mercury reductions, but rather from highly imprecise estimates of mortality reductions from decreasing emissions of fine particulate matter, or $PM_{2.5}$. Importantly, these estimates are not based on an evaluation of all available relevant science; rather, EPA

relied on two observational epidemiology studies conducted when air pollution levels were generally above current standards.

Epidemiology studies investigate statistical associations or correlations between estimated levels of air pollutants and health outcomes in human populations. The two studies on which EPA relied reported statistical associations between $PM_{2.5}$ reductions and health benefits and assumed a causal relationship, but dozens of other epidemiology studies are available and many report no such correlations. The fact that EPA only considered studies that suggested an association means that it conducted a biased assessment of the available data.

Even if it were appropriate to rely only on these two studies, just because two factors are correlated does not mean that one caused the other; study outcomes can depend on many factors besides pollution. For example, health risk factors – such as smoking, exercise, and diet – may have contributed to the increased mortality some studies attributed to $PM_{2.5}$. In addition, most epidemiology studies, including the two on which EPA relied, estimated personal exposure from monitors at central sites, even though most people spend a majority of their time indoors. These monitors do not accurately capture daily variations in $PM_{2.5}$ concentrations or composition that may differ from what is experienced by individuals, particularly indoors. This leads to inaccurate results in epidemiology analyses.

Finally, in addition to ignoring much of the epidemiology evidence, EPA did not consider other lines of evidence in its benefits estimations. Experimental studies have demonstrated that the physiological impacts of inhaling $PM_{2.5}$ are only observed when very high doses overwhelm the

lungs' natural defense mechanisms. In other words, the body's natural defenses can effectively deal with a certain level of $PM_{2.5}$. Above that level, called a threshold, additional $PM_{2.5}$ can perturb normal function. Indeed, some level of $PM_{2.5}$ in ambient air is unavoidable and has been present on earth for eons, but humans have evolved the means to cope with these exposures without major health consequences.

Despite this, EPA assumed that there is no level of $PM_{2.5}$ below which health effects, including mortality, would not be observed. Although EPA acknowledged that the benefits estimates would be significantly overestimated if a threshold was incorporated in its analyses, it nonetheless calculated benefits without one. If a threshold were accounted for, mortality estimates would be much less – and could be zero.

In conclusion, the largest benefits from the Mercury and Air Toxics Standards are derived not from reducing mercury, but from reducing $PM_{2.5}$. Despite the vast array of peer-reviewed scientific literature on the topic, EPA based its calculations on only two epidemiology studies. These two studies had several methodological limitations, including the inability to assess alternative causes of the observed health effects and the reliance on central monitors to estimate personal exposures. These studies were not consistent with many epidemiology studies indicating no correlation between reducing $PM_{2.5}$ and health benefits, nor experimental studies indicating an exposure threshold below which $PM_{2.5}$ is not likely to overwhelm the body's natural defenses.

All of these factors indicate that the benefits estimates from the Mercury and Air Toxics Standards are grossly inflated and not realistic. Because there is arguably very limited evidence that these standards would reduce the disease burden more than pollution standards already in place, resources should be used towards other measures that would more clearly benefit society.

Thank you again for the opportunity to testify today and I look forward to answering your questions.

Biographical Summary

Julie E. Goodman, Ph.D., DABT
Principal

Dr. Goodman is an expert in toxicology, epidemiology, and assessing human health risks from chemicals in consumer products and the environment. Her primary responsibilities at Gradient include the design, oversight, analysis, and interpretation of epidemiology studies as well as the evaluation of chemical toxicology data, apparent disease clusters, and chemical exposures. Before joining Gradient, Dr. Goodman was a Cancer Prevention Fellow at the National Cancer Institute. Dr. Goodman has authored original research articles, review articles, and book chapters on a wide variety of topics related to epidemiology and toxicology, including weight-of-evidence analyses of several chemicals. She also has presented scientific findings and analyses to community groups and regulatory and legislative bodies. She is currently an adjunct faculty member in the Department of Epidemiology at the Harvard School of Public Health.

Representative Projects

Cancer Cluster Analysis: At the request of a municipality and in response to citizens' concerns, investigated whether there was an increased incidence rate of cancer in residents living near a municipal landfill. Communicated findings to city officials and residents at public meetings.

Cross-Sectional Study: Critically reviewed cancer and noncancer trichloroethylene and perchloroethylene toxicity data. Conducted quantitative analysis of exposure to these solvents in groundwater *via* ingestion and showering. Determined whether health effects in an allegedly exposed community were comparable to those in communities with no known solvent exposures based on questionnaire data.

Efficacy and Toxicity Analysis: For a pharmaceutical company whose patent was being challenged, performed an independent analysis of efficacy and toxicity data to determine whether claims in the patent could be challenged.

Regulatory Comment: Provided written and oral comments to the Clean Air Scientific Advisory Committee (CASAC) on clinical and epidemiological studies and their bearing on US EPA's development of National Ambient Air Quality Standards (NAAQS) for ozone, particulate matter, nitrogen oxides, and sulfur oxides.

Weight-of-Evidence Analysis: Conducted a comprehensive critical weight-of-evidence review of studies bearing on the ability of very low exposures to bisphenol A to affect reproduction and development *via* endocrine disruption. Testified before several state legislative committees regarding potential restrictions on bisphenol A.

Benchmark Dose Calculations: Analyzed US EPA's use of the lower confidence limit on the BMD_1 ($BMDL_1$) to determine a point of departure for cancer risk of dimethylarsenic acid in humans in a white paper submitted to US EPA.

Product Safety Analysis: Determined whether a toxicological evaluation of a toy was sufficient for determining children's health risks. Conducted an independent analysis of potential routes of exposure to and toxicity of a chemical found in the toy.

Meta-analysis: Conducted meta-analyses and meta-regressions of airway hyper-responsiveness data from clinical studies of asthmatic volunteers exposed to NO_2 while exercising or at rest.

Practice Areas & Expertise

- Epidemiology
- Toxicology
- Occupational Exposures
- Product Safety
- Carcinogenesis
- Risk Assessment

Education

Ph.D., Toxicology, Johns Hopkins University

Sc.M., Epidemiology, Johns Hopkins University

S.B., Environmental Engineering, Massachusetts Institute of Technology

Diplomate of the American Board of Toxicology

Selected Publications

Rhomberg, LR; Bailey, LA; Goodman, JE; Hamade, AK; Mayfield, DB. 2011. "Is exposure to formaldehyde in air causally associated with leukemia? – A hypothesis-based weight-of-evidence analysis." *Crit. Rev. Toxicol.* (In Press).

Goodman, JE; Dodge, DG; Bailey, LA. 2010. "A framework for assessing causality and adverse effects in humans with a case study of sulfur dioxide." *Reg. Tox. Pharmacol.* 58:308-322.

Goodman, JE; Kerper, LE; Petito Boyce, C; Prueitt, RL; Rhomberg, LR. 2010. "Weight-of-evidence analysis of human exposures to dioxins and dioxin-like compounds and thyroid hormone levels during early development." *Reg. Tox. Pharmacol.* 58(1):79-99.

Goodman, JE; Nascarella, MA; Valberg, PA. 2009. "Ionizing radiation: A risk factor for mesothelioma." *Cancer Causes and Control.* 20:1237-1254.

Goodman, JE; Prueitt, RL; Dodge, DG; Thakali, S. 2009. "Carcinogenicity assessment of water-soluble nickel compounds." *Crit. Rev. in Toxicol.* 39(5):365-417.

Gradient

20 University Road
Cambridge, MA 02138
Phone (617) 395-5000
Fax (617) 395-5001

jgoodman@gradientcorp.com
www.gradientcorp.com

Appendix to the Testimony of Julie E. Goodman, Ph.D., DABT

Regarding
EPA's Assessment of Health Benefits
Associated with $PM_{2.5}$ Reductions
for the Final Mercury and Air Toxics Standards

Prepared for

The American Energy Initiative Hearing

Congress of the United States
House of Representatives
Committee on Energy and Commerce
Subcommittee on Energy and Power
2125 Rayburn House Office Building
Washington, DC 20515-6115

February 8, 2012

Estimation of Health Benefits from Reductions of PM

The US Environmental Protection Agency (EPA) recently released "Benefits and Costs of the Clean Air Act Report from 1990 to 2020" (US EPA, 2011a) and several associated documents that present the underlying methodology (Industrial Economics, Inc. (IEc), 2006, 2010, 2011). This report, also called the "Second Prospective Study," is the third in a series of EPA studies that evaluated programs related to the implementation of the Clean Air Act (CAA) and its 1990 Amendments (CAAA).

Approximately 90% of the economic benefits reported in the Second Prospective Study relate to reductions in mortality associated with particulate matter (PM) and ozone (O_3); the remaining benefits are divided between reductions in illness (morbidity) and visibility improvements. The majority of the issues discussed below are also relevant to analyses conducted for the Mercury and Air Toxics Standards.

The likely largest source of uncertainty in the CAAA benefits estimation is the choice of the concentration-response function[1] (CRF) that relates the reduction in $PM_{2.5}$ air concentrations to reductions in adverse health outcomes. Underlying this choice is the assumption that statistically significant associations reported in the epidemiology literature are causal. Although EPA acknowledged that "[i]f the PM/mortality relationship is not causal, it would lead to a significant overestimation of net benefits" (US EPA, 2011a, Table 5-11), it did not consider any non-causal scenarios. There are many epidemiology studies that find no association between PM and mortality.

EPA relied heavily on the epidemiology literature in its evaluation of the health impacts from air pollutants and in selecting appropriate CRFs, even though studies report mixed results in the case of PM-associated mortality. While the two studies on which EPA relied report positive statistically significant effects (*e.g.*, Pope *et al*., 2002 and Laden *et al*., 2006), other studies show no effect (*e.g*., Beelen *et al*., 2008; Brunekreef *et al.*, 2009; Enstrom, 2005; McDonnell *et al*., 2000; Lipfert *et al*., 2006; Zeger *et al*., 2008). EPA placed no weight on these latter studies, and thus did not consider a possible null or no-effect association in the Second Prospective Study.

The first study on which EPA relied to quantify the deaths avoided from $PM_{2.5}$ is a re-analysis of the American Cancer Society (ACS) cohort by Pope *et al*. (2002); the second is a re-analysis of the Harvard Six Cities (HSC) Study by Laden *et al*. (2006). Although these studies have undergone a limited

[1] The concentration-response function describes the change in effect on an organism caused by differing levels of exposure to a stressor after a certain exposure time.

amount of reanalysis, there are remaining limitations that make them unreliable in a quantitative analyses, particularly if considered in isolation from the results from other epidemiology studies.

These two studies reported different mortality estimates. Pope *et al*. (2002) found a 0.6% increase in all-cause mortality per $\mu g/m^3$ of $PM_{2.5}$, while Laden *et al*. (2006) found a 1.5% increase in all-cause mortality. EPA used the mid-point between these two estimates in its benefits analysis (*i.e.*, 1% per $\mu g/m^3$ of $PM_{2.5}$), and gave two bases for its choice: Its assumption that the ACS study underestimated responses because this cohort had a greater percentage of white, educated, higher income participants that are less representative of the susceptible population compared to the HSC study; and its assumption that the ACS study had more exposure measurement error because it relied on a single central monitor in each large city compared to the HSC study, which used monitors that were specifically located for the study. As discussed below in Sections 1.2 and 1.3, neither of these reasons is scientifically sound and raise questions about the magnitude of the estimated mortality effects.

Additional sources of uncertainty discussed in greater detail in Section 1 include the reliability of statistical models used and how effectively the models can control for confounding factors. In Section 2, significant uncertainty in the shape of the CRF is discussed amid mounting evidence that a threshold for PM-related effects exists. In Section 3, EPA's assumption that all PM is similarly toxic is discussed.

1 Uncertainty in the Magnitude of the Mortality Estimate for Particulate Matter

Not only is the question of causality unresolved, but questions remain as to the magnitude of the effects reported in the epidemiology literature. In the Second Prospective Study, EPA relies on two studies as the basis for the CRFs for PM mortality (Pope *et al*, 2002; Laden *et al.*, 2006), although there are a number of other long-term mortality studies that should have been considered. Several studies report no association between PM and mortality, yet EPA does not acknowledge them. EPA's justification for inclusion of the HSC and ACS studies is flawed, and a number of uncertainties in the epidemiology findings raise questions about their use in quantitative benefits assessments. Some of the key uncertainties include exposure measurement error, confounding, and model specification.

1.1 Choice of Concentration-Response Function

EPA limited its choice of CRFs to those from only two studies, not considering the full range of studies available. Importantly, several recent long-term mortality studies have reported no association between $PM_{2.5}$ and mortality, and EPA does not include the possibility of no causal association between PM and mortality in its estimated benefits analyses.

For example, analyses of a large Netherlands Cohort (the NLCS-Air) have reported consistently null results in investigations of PM-related mortality (*e.g.*, Beelen *et al*., 2008; Brunekreef *et al.*, 2009). Similarly, McDonnell *et al*. (2000) reported no association between $PM_{2.5}$ concentrations and mortality in a large cohort of Seven Day Adventists in California. In another study, Zeger *et al*. (2008) found a lack of association between $PM_{2.5}$ concentrations and mortality for the western US regions, whereas a statistically significant association was reported for the eastern and central regions of the country. Similarly, Lipfert *et al*. (2006) reported a weak association between mortality and $PM_{2.5}$ in single-pollutant models, but no association was noted when they included traffic density in the analyses of a large veterans cohort. Also, Enstrom (2005) reported no association between fine PM and chronic mortality in elderly Californians.

Instead of considering the full range of potential CRFs from the available epidemiology literature, including those that show no or "beneficial" effects of PM, EPA relied on expert elicitation to support its choice of a CRF, asking 12 experts to propose mortality estimate distributions associated with long-term PM exposures (Roman *et al*., 2008; IEc, 2006).

EPA used expert judgment elicitation as a means of capturing the uncertainty in the CRF. The use of experts to attain this information opens the question of bias in the choice of expert judgments, particularly since the group was not a random sample of experts representing the range of scientific opinions on the subject. For example, six of the 12 experts were co-authors of the ACS and HSC studies, which EPA ultimately relied on to quantify PM mortality. Also, the opinions of experts should not be a substitute for empirical data. In fact, as discussed by Roman *et al*. (2008), one of the challenges in the elicitation study was how to reconcile expert opinion on the likelihood of a causal relationship with the CRF function uncertainty distribution. For example, one expert opined that the likelihood of a causal association was 35%, yet his uncertainty distribution did not include a 0% decrease in mortality per $1\mu g/m^3$ $PM_{2.5}$.

Skepticism that expert elicitation is appropriate for use in quantitative risk assessment is shared by the NRC Committee on Improving Risk Assessment Approaches (CIRAA), commissioned by EPA to provide advice on improving its risk assessment process (NRC, 2009). This committee was concerned with both the methodology and use of expert elicitation.

Regardless of adequacy of expert elicitation, results of the EPA expert elicitation distributions varied widely by expert, although all were positive. Overall, eight out of 12 experts estimated a PM-associated mortality that was lower than the primary estimate that EPA used (mean of 27% over a 1-80% range). This is consistent with benefits from the CAAA being overestimated.

In summary, EPA did not consider the available epidemiology research fully in developing the CRF for use in its quantitative assessment of mortality reductions associated with reduced $PM_{2.5}$ levels. In addition, because it did not consider a lower bound to the estimates inclusive of a null or non-causal association between PM and mortality, the estimates provided in its Second Prospective Study are likely biased high with significant uncertainties understated.

1.2 Effects on Susceptible Population Groups

The ACS study by Pope *et al*. (2002) included a cohort of over 1 million adults in over 50 US cities, but was a more homogenous population than the general US population. EPA concluded that the authors likely underestimated any mortality effects because the study did not sufficiently represent potentially susceptible population groups, such as people with a lower socioeconomic status (SES). EPA cited the re-analysis of the ACS study conducted by Krewski *et al*. (2000) as evidence of potential effect modification based on SES.[2] There is little evidence to support that socioeconomic factors modify mortality estimates as the data regarding effects of SES on PM mortality associations are inconclusive at best. EPA actually noted in the Second Prospective Study that the direction of the bias associated with this source of uncertainty cannot be determined based on available data (US EPA, 2011a, Table 5-11).

As part of a sensitivity analysis, Krewski *et al*. (2000) identified potentially "susceptible" subgroups and conducted analyses for each subgroup. The only modifying factor that was found to have a significant effect on PM-associated mortality was education (chosen as a surrogate of SES). In the ACS cohort, Krewski *et al*. (2000) found larger risks of mortality in a subpopulation of people with less than a

[2] An effect modifier is a factor that results in a change in the magnitude of an association between an exposure and an outcome when data are stratified by that factor (Last, 2001).

high school education than in the full cohort. Conflicting results were reported in the most recent analysis of the ACS cohort, which extended the follow-up time to 18 years, from 1982-2000 (Krewski *et al*., 2009). As in the previous analyses, the most current evaluation featured sensitivity analyses that assessed effect modification by education. For this follow-up, however, a trend of effect modification by education was more difficult to discern and for some health outcomes (*e.g.*, ischemic heart disease), there was a reverse trend such that greater risks were observed for the more educated. It is unknown whether the SES risk gradient observed indicates a higher risk in those with lower SES, or alternatively, as Krewski *et al.* (2009) reported, that there may be inadequate control for socioeconomic factors in the study.

Few studies are available that specifically address SES modification by $PM_{2.5}$ exposures, but several studies have assessed the modifying effects of other PM fractions. Overall, the evidence is mixed. Laurent *et al.* (2007) recently reviewed epidemiology studies of the interaction between SES and air pollution-related mortality (including PM). The authors were not able to make formal comparisons between studies due to the large variety of SES indicators used across the studies. One important finding was that no effect modification by SES was found in studies that used SES indicators at coarse geographic resolutions (city or county level), whereas mixed results were reported for studies that used SES measures at finer geographic resolutions. Overall, the authors noted that there is not enough information to conclude that SES modifies the relationship between air pollution and mortality outcomes.

Although each community in the HSC cohort included a more heterogeneous population than the ACS cohort, the study was much smaller and limited to six cities in the midwestern and northeastern US that are unlikely to be representative of the overall US population or the mix of air pollutants and other factors across the US.

Overall, EPA provided weak justification for focusing on the much higher reported mortality estimates from the Laden *et al*. (2006) analysis, as the literature is not supportive of a "larger" mortality effect from $PM_{2.5}$ exposures in lower SES populations. In addition, EPA does not provide justification for not considering the full range of possible CRF functions available in the literature, which are not limited to the results from these two studies.

1.3 Exposure Measurement Error

As EPA notes, the Pope *et al*. (2002) and Laden *et al*. (2006) studies are also limited in that both studies had to estimate $PM_{2.5}$ concentrations for a large part of the follow-up period (1980s & 1990s) because there were no $PM_{2.5}$ measurements available. Even if these data were available for all years, these studies relied on central monitors to estimate personal exposure, which led to exposure measurement error.

In the ACS study, researchers used average $PM_{2.5}$ concentrations based on the early and later study periods, whereas, in the HSC study, Laden *et al*. (2006) used city-specific regression equations based on extinction coefficients, collected PM_{10} concentrations from monitors within 80 km of study subjects' homes, and indicators for season to estimate $PM_{2.5}$ concentrations for years when measurements were not available. This also introduced uncertainty into the association between $PM_{2.5}$ and mortality. The amount and direction of the bias in both studies are uncertain, but likely overestimated risk associated with exposures to $PM_{2.5}$ (Rhomberg *et al*., 2011).

Exposure assessment studies have shown that central site data do not adequately represent personal exposure, in part because most people spend a large portion of their time indoors (Lioy *et al*., 1990; Mage and Buckley, 1995; Janssen *et al*., 1997, 1998; Ozkaynak *et al*., 1996; Dominici *et al*., 2003). Exposure measurement error occurs because central-site monitors may not accurately capture population mobility, the uneven distribution of PM exposure attributable to local sources, pollution patterns that can be affected by terrain features and weather, and daily variations in PM concentrations or composition that may differ from variations experienced by individuals. These factors may bias the results of an epidemiology analysis in either direction, and are particularly relevant for long-term studies for which these factors likely also vary over time. The direction and magnitude of the bias depends on the type of measurement error and spatial variability of air pollutant concentrations is likely to result in effects being overestimated (Goldman *et al*., 2011).

Exposure measurement error also affects the interpretation of the CRF for air pollution effects. EPA has often dismissed this important source of uncertainty assuming that the bias is likely to be towards the null. In Second Prospective Study, EPA indeed stated that this bias likely underestimated the benefits (US EPA, 2011a, Table 5-11). Recent studies have shown that this bias can be in either direction but the type of bias typically associated with spatially variable pollutants usually overestimates effects.

1.4 Confounding Bias

A large source of uncertainty that is common to all air pollution epidemiology studies is confounding. A confounder is a factor associated with both the exposure and the health outcome, but is not causal. For example, individual risk factors (*e.g.*, smoking, diet, *etc.*) may contribute to or even fully explain the deaths attributed to PM. The main challenge is the large number of potential confounders which include co-pollutants, temporal trends, individual factors, and meteorological factors.

The study by Pope *et al.* (2002) analyzed potential confounding factors. The researchers tested confounding by smoking, education, body mass index (BMI), diet, alcohol consumption, and occupational/other exposures. Although mortality risk reductions were observed when controlling for these individual factors, the reductions were not statistically significant. While it is plausible that these factors did not play a role in the observed association, it is also likely that they were not accurately estimated in the study because these risk factors were assessed only at the time of enrollment, nearly thirty years ago, and changes in these risk factors were not assessed during follow-up. Furthermore, the SES factors in this study were collected using a self-administered questionnaire, an approach that is well known to result in under-reporting of key potential confounding risk factors for mortality (*e.g.* smoking).

In the Pope *et al*. (2002) study, spatial confounding (effects that may be due to regional or other spatial differences across cities) was explored by applying complex statistical modeling (*i.e.*, random effects models). The results indicated that for all-cause mortality, effect estimates were reduced to statistical insignificance when regional differences were included in the model. This indicates that confounding was likely not fully accounted for in the study.

In addition, Pope *et al*. (2002) assessed mortality associations with alternative PM metrics [*e.g.*, coarse particulate matter (PM_{10}) and total suspended particles (TSP)], sulfates, and various gaseous pollutants (*e.g.*, SO_2, NO_2, CO, and O_3). The mortality estimates associated with sulfates and SO_2 were of the same magnitude as the $PM_{2.5}$-related estimates, but the researchers found no association for other PM metrics and no association with O_3. Interestingly, the authors did not assess potential confounding of the $PM_{2.5}$ mortality association by SO_2 and sulfate in two-pollutant models, even though a reanalysis of the original study indicated these pollutants significantly confounded the PM mortality associations (*e.g.*, Krewski *et al*., 2000). This is a very critical omission. The ambient levels of SO_2 have decreased markedly since the initiation of the ACS study. It is possible that at the current levels of SO_2, researchers would find no significant association between ambient PM and mortality.

The bias associated with confounding effects is particularly difficult to address in epidemiology studies because it is often difficult to account for all potential confounding factors. In PM mortality studies there is evidence that co-pollutants can confound the PM mortality association, particularly strongly correlated pollutants such as SO_2. Even if potential confounders are accounted for in studies, there may still be issues of how well the confounding variables are measured and, as with the Pope *et al*., (2002) study, whether confounders were re-evaluated over the follow up study period. The issue of confounding relates to both the assumption of causality, where another factor may actually be the causal agent, and to the magnitude of the association, where a co-factor may account for some of the observed risk. In the Second Prospective Study, EPA did not address the potential bias associated with confounding either quantitatively or qualitatively.

1.5 Model Selection Bias

A remaining large source of uncertainty in the PM mortality association involves how different statistical models impact epidemiology findings. To address this question, researchers conduct extensive sensitivity analyses, including tests of the effects of various model assumptions (*e.g.*, lags and smoothing functions for time trends), to assess the impacts on mortality estimates. There have been questions raised on the appropriateness of the standard Cox Proportional Hazards Model that was used by the two studies EPA relied on for the PM CRFs (Pope *et al*., 2002; Laden *et al*., 2006).

A risk estimate is dependent on the statistical model from which it is calculated. If a model is based on assumptions that are not met, risk estimates are likely biased. For example, Moolgavkar (2005) notes that the assumptions of the Cox proportional hazards model are violated in ecological studies of pollution health effects. This is likely for several long-term $PM_{2.5}$ exposure studies, including the study by Laden *et al*. (2006). As Abrahamowicz *et al*. (2003) noted:

> [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.

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 ACS, which included 50 cities with $PM_{2.5}$ data, and 151 cities

with sulfate data. They found a statistically significant deviation from the traditional linearity assumption for both $PM_{2.5}$ and sulfate. They also found that risk estimates for both $PM_{2.5}$ and sulfate differed from those based on models using the traditional assumptions, with $PM_{2.5}$ risks inflated at low doses, and sulfate showing a threshold. These results illustrate that the Cox PH models give inaccurate risk estimates, particularly at low doses.

Koop and Tole (2004) also emphasized that by neglecting the important issue of model uncertainty, or the choice of a specific model among the many options assessors have, "most studies overstate confidence in their chosen model and underestimate the evidence from other models," and can result in "uncertain and inaccurate results." Furthermore, the authors found that when model uncertainty is incorporated into the estimation of air pollution effects, it is so large that the plausibility of effects become questionable. These authors argue that such estimates not be used in policy decision-making, which excludes their use in quantifying impacts of regulations.

In summary, recently conducted analyses to test how model choice impacts mortality estimates find a significant impact on results for one of the most commonly used models for long-term mortality effects analyses, the Cox-PH mode. Model uncertainty has generally not been incorporated in the estimates of air pollution effects and if it is considered, it would likely result in many non-statistically significant results. As with confounding bias, EPA does not address the impact of model uncertainty in its selected CRF function.

2 Uncertainty in the Shape of the CRF for PM Mortality

As noted above, questions remain regarding the shape of the CRF. EPA assumed that the PM-mortality relationship is linear at low concentrations with mortality directly proportional to the ambient particle concentration. The uncertainty of the linear coefficient describing the relationship is considered, but the possibility that the function is nonlinear is not given the same consideration. EPA qualitatively discussed this potentially large source of uncertainty, noting that the bias would overestimate the benefits, but concluded that the effects would be minor. The sensitivity analyses conducted in the First Prospective Study, however, demonstrated that considering a threshold had significant effects on mortality estimates. Several studies provide evidence that the PM-mortality association is non-linear and that a threshold exists. For example, Smith *et al*. (2000) reported PM mortality thresholds at 20-25 μg/m^3. As shown in Figure 2.1, based on the EPA sensitivity analysis, a threshold at 20 μg/m^3 would decrease avoided deaths from ~20,000 to 5,000 or fewer (US EPA, 1999).

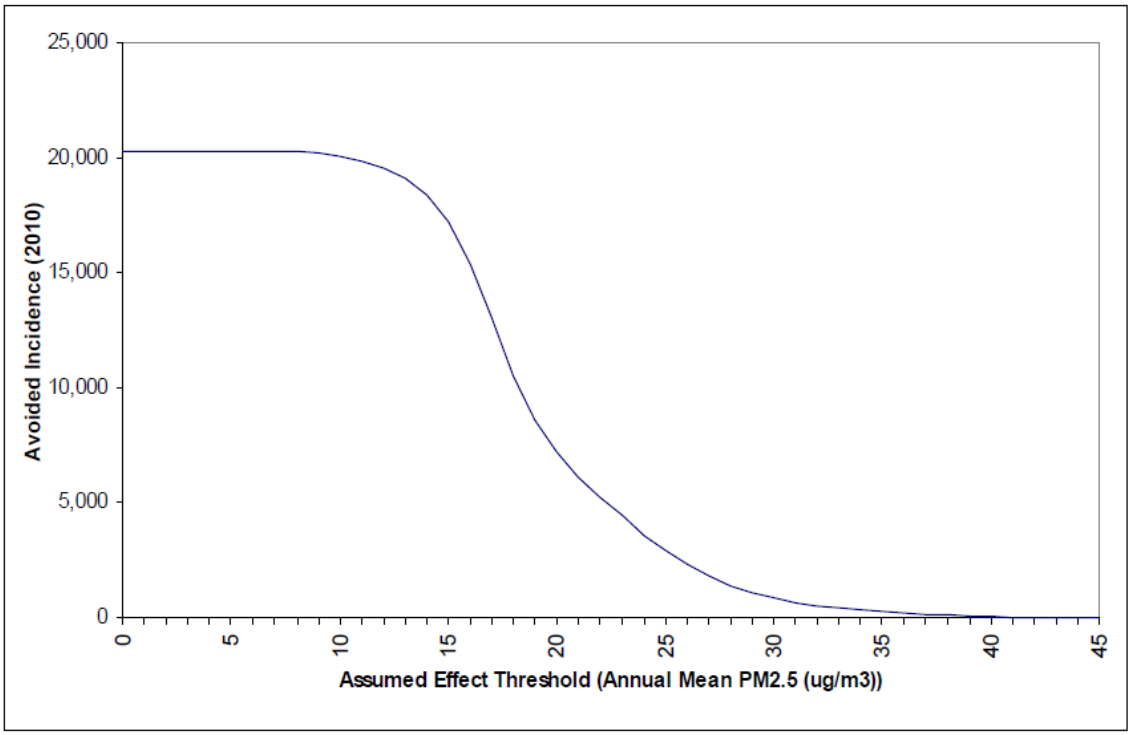

Figure 2.1 Long-term Mortality Incidence Avoided Assuming Different $PM_{2.5}$ Thresholds. Based on the CRF from Pope *et al*. (1995). Source: US EPA, 1999, Figure D-2.

A threshold for PM health effects is supported by toxicological, occupational, and human exposure evidence. Toxicological studies demonstrate that the physiological impact and biological mechanism of inhaled PM effects comes from overwhelming the natural defense mechanisms from the mass of particles deposited locally onto tissues (*e.g.*, Oberdorster, 1996, 2002; Pauluhn, 2011; Valberg *et al*., 2009). Therefore, one would expect to see thresholds and/or nonlinear behavior with higher doses. Indeed, animal studies using carbon black and titanium dioxide (TiO_2) particles show that a threshold for PM-related effects exists (Oberdorster, 1996, 2002). Furthermore, the EPA Health Assessment Document for Diesel Exhaust (US EPA, 2002) reports a no observable adverse effect level (NOAEL) for chronic diesel exhaust particulate (DEP) exposures of 460 $\mu g/m^3$. This NOAEL is based on EPA's comprehensive review of the large numbers of laboratory-animal studies with exposures up to high levels of diesel exhaust (100-7,800 $\mu g/m^3$). Similarly, the development of occupational standard threshold limit values show that other government agencies have been able to derive threshold level of effects for many types of particles (Oller and Oberdörster, 2010). Lastly, human exposure studies using DEP suggest a threshold for inflammatory responses (*e.g.*, Mudway *et al*., 2004; Behndig *et al*., 2006; Peretz *et al*., 2008a,b).

Assuming a linear relationship has significant impacts on health effects benefits estimates because, when a linear function is used to describe health impact for an effect that is truly nonlinear with exposure, then the effect on health is overestimated at lower concentrations and may be (depending on the range of concentrations) underestimated at high concentrations. This is because the change in estimated effect brought about by a reduction in exposure levels depends heavily on how those reductions are distributed over the range of exposure (Rhomberg *et al*., 2011).

For example, benefits of a control program that knocks down the upper end of the exposure range, but leaves the lower end largely unchanged will tend to be undervalued because the assumed linear function fails to attribute most of the original mortality impact to high-end exposures. Further, this method fails to note that most of the exposure reduction occurs at the high end, where it is most effective.

In contrast, a program that generally lowers all exposure levels but does not disproportionally lower high-end exposures will tend to be overvalued, because it ascribes illusory benefits to the reductions of the already low exposures experienced by much of the population. Indeed, because most of the population exposure occurs at the lower parts of the distribution of exposures even small overestimates of the benefits can, when collected over such a large fraction of the population, dominate the population benefit.

Observed linear relationships between PM exposures and mortality may be artificial due to exposure measurement error. That is, in addition to affecting the magnitude of the effect estimate, exposure measurement error also influences the **s**hape of the CRF. This is because some individuals in the population have greater exposures than others for any given central-site ambient concentration. This will artificially flatten apparently linear CRFs and make concentration-related effects (even those that are truly threshold in nature) look linear, masking what may in fact be a steeper curve (Brauer *et al*., 2002; Rhomberg *et al*., 2011).

For example, Meng *et al*. (2005) hypothesized that biases arise in $PM_{2.5}$–health effects associations because of seasonal variations in infiltration behavior. Their data showed that seasonal differences in infiltration behavior coincide with fluctuations in ambient PM concentrations and vary with location. In particular, they found that during the summer, when $PM_{2.5}$ concentrations are generally higher, there was an increase in infiltration factors in New Jersey homes from opening of windows for ventilation, whereas in Texas there was a reduction in infiltration factors because of the use of air conditioners. The researchers concluded that exposure measurement error from differences in infiltration behavior bias health estimates in chronic studies. The magnitude of the error can differ between communities and differentially impact personal-ambient relationships – *e.g.*, mean ambient $PM_{2.5}$ concentrations could be higher in City A *vs.* City B, but due to differences in particle infiltration behavior in the two cities, mean exposures to ambient $PM_{2.5}$ could be reversed. Dominici *et al*. (2002) also reported nonlinear C-R curves when analyzing data at the regional level and noted that nonlinearities are likely averaged out in multi-city studies that present national CRFs.

In conclusion, EPA assumed a linear relationship in its calculation of health impacts from exposure to PM. Evidence is growing in the epidemiology literature that this relationship is in fact nonlinear, and that factors such as exposure measurement error and pooling multi-city effect estimates lead to the appearance of a linear relationship. A threshold for PM effects is also supported by toxicological, occupational, and human chamber studies. A threshold was assumed in the sensitivity analysis conducted for the First Prospective Study, showing much lower mortality incidence when a threshold is assumed.

3 Differential Toxicity of PM Size fractions and PM Components

An additional important source of uncertainty in the CRFs is the regional and seasonal heterogeneity in $PM_{2.5}$ concentrations, population characteristics, and risk estimates that introduce additional bias to overall effect estimates in epidemiology studies (US EPA, 2010, 2011b). For example, in multi-city studies that employ a common model specification, risk estimates may be biased due to differences in $PM_{2.5}$ sources, $PM_{2.5}$ composition, $PM_{2.5}$ concentrations, the adequacy of central monitors to measure personal exposures, and/or population characteristics (*e.g.*, personal behaviors or susceptibilities). Researchers have found significant differences in effect estimates across cities and regions that are unexplained despite recent efforts to evaluate modifying effects that could account for these differences (US EPA, 2011b).

The $PM_{2.5}$ NAAQS makes no distinction between components of $PM_{2.5}$, treating all $PM_{2.5}$ as equally toxic. However, the spatial, temporal, and toxicological composition of $PM_{2.5}$ can vary greatly. The uncertainty associated with differential toxicity of $PM_{2.5}$ components can be significant, as discussed in the IEc uncertainty analyses report (IEc, 2010). Control strategies that reduce specific $PM_{2.5}$ components also affect other components, adding to the complexity of the issue.

For example, regulations that specifically reduces sulfates and nitrates also affect ammonia. In certain parts of the country, these three $PM_{2.5}$ components make up about 40-50% of the PM mass, mostly derived from gas to aerosol conversion from large point sources (such as utilities and industrial combustors) (Green *et al*., 2002). There is no evidence either from human exposure studies or animal studies, however, to suggest that sulfates, nitrates, or ammonia at current ambient levels are associated with mortality or morbidity outcomes (Green *et al*., 2002; Utell *et al*., 1983; US EPA, 1996).[3] Therefore, if controls are focused on particulate components that are highly unlikely to contribute to mortality, and if these PM reductions are counted as contributors to the avoided mortality, then these controlled benefits would be exaggerated and misleading. In the Second Prospective Study, because EPA assumed that all PM is of equal toxicity, the benefits estimates are thus likely biased high.

Although the particulate composition and differential toxicity issue is currently being investigated as noted in the Uncertainty report (IEc, 2010), there is no clear resolution. This issue remains a

[3] Airborne sulfate is widely used in medicine. It is a common ingredient in bronchodilators used to treat asthma. If fact, one puff of an albuterol sulfate inhaler delivers sulfate at a concentration of about 10,000 µg of sulfate per m^3 of inhaled air (Green *et al*., 2002) and is not only considered safe, but beneficial.

potentially significant source of uncertainty in both the assumption of a causal relationship between PM and health effects (particularly mortality) and if a causal relationship exists at low levels, in the magnitude of these effects.

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.

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.

Behndig, AF; Mudway, IS; Brown, JL; Stenfors, N; Helleday, R; Duggan, ST; Wilson, SJ; Boman, C; Cassee, FR; Frew, AJ; Kelly, FJ; Sandström, T; Blomberg, A. 2006. "Airway antioxidant and inflammatory responses to diesel exhaust exposure in healthy humans." *Eur. Respir. J.* 27(2):359-365.

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.

Brunekreef, B; Beelen, R; Hoek, G; Schouten, L; Bausch-Goldbohm, S; Fischer, P; Armstrong, B; Hughes, E; Jerrett, M; van den Brandt, P. 2009. "Effects of long-term exposure to traffic-related air pollution on respiratory and cardiovascular mortality in the Netherlands: The NLCS-AIR Study." Health Effects Institute, Boston, MA.

Dominici, F; Daniels, M; Zeger, SL; Samet, JM. 2002. "Air pollution and mortality: Estimating regional and national dose-response relationships." *J. Amer. Stat. Assoc.* 97(457):100-111

Dominici, F; McDermott, A; Zeger, SL; Samet, JM. 2003. "National maps of the effects of particulate matter on mortality: Exploring geographical variation." *Environ. Health Perspect.* 111(1):39-44.

Enstrom, JE 2005. "Fine particulate air pollution and total mortality among elderly Californians, 1973-2002." *Inhal. Toxicol.* 17:803-816.

Goldman, GT; Mulholland, JA; Russell, AG; Strickland, MJ; Klein, M; Waller, LA; Tolbert, PE. 2011. "Impact of exposure measurement error in air pollution epidemiology: Effect of error type in time-series studies." *Environ. Health* 10(1):61.

Green, LC; Crouch, EAC; Ames, MR; Lash, TL. 2002. "What's wrong with the national ambient air quality standard (NAAQS) for fine particulate matter (PM2.5)?" *Regul. Toxicol. Pharmacol.* 35:327-337.

Industrial Economics, Inc. (IEc). 2006. "An Expert Judgment Assessment of the Mortality Impact of Changes in Ambient Fine Particulate Matter in the U.S.: Supporting Information." Report to US EPA, Office of Air Quality Planning and Standards. 413p.

Industrial Economics, Inc. (IEc). 2010. "Uncertainty Analysis to Support the Second Section 812 Benefit-Cost Analysis of the Clean Air Act (Draft)." Report to US EPA, Office of Air and Radiation. 200p., April.

Industrial Economics, Inc. (IEc). 2011. "Health and Welfare Benefits Analyses to Support the Second Section 812 Benefit-Cost Analysis of the Clean Air Act (Final)." Report to US EPA, Office of Air and Radiation. 164p., February.

Janssen, NA; Hoek, G; Brunekreef, B; Harssema, H; Mensink, I; Zuidhof, A. 1998. "Personal sampling of particles in adults: relation among personal, indoor, and outdoor air concentrations." *Am J Epidemiol.* 147(6):537-547.

Janssen, NA; Hoek, G; Harssema, H; Brunekreef, B. 1997. "Childhood exposure to PM_{10}: Relation between personal, classroom, and outdoor concentrations." *Occup Environ Med.* 54(12):888-894.

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

Krewski, D; Burnett, RT; Goldberg, M; Hoover, K; Siemiatycki, J; Jerrett, M; Abrahamowicz, M; White, WH; et al. 2000. "Reanalysis of the Harvard Six Cities Study and the American Cancer Society Study of Particulate Air Pollution and Mortality." Health Effects Institute (Cambridge, MA). 9p., August.

Krewski, D; Jerrett, M; Burnett, RT; Ma, R; Hughes, E; Shi, Y; Turner, MC; Pope, A; Thurston, G; Calle, EE; Thun, MJ. 2009. "Extended Follow-Up and Spatial Analysis of the American Cancer Society Study Linking Particulate Air Pollution and Mortality." HEI Research Report 140, Health Effects Institute (Cambridge, MA). May.

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.

Last, JM. 2001. *A Dictionary of Epidemiology, 4th ed.* Oxford University Press, New York.

Laurent, O; Bard, D; Filleul, L; Segala, C. 2007. "Effect of socioeconomic status on the relationship between atmospheric pollution and mortality." *J. Epidemiol. Community Health* 61:665-675.

Lioy, PJ; Waldman, JM; Buckley, T; Butler, J; Pietarinen, C. 1990. "The personal, indoor and outdoor concentrations of PM-10 measured in an industrial community during the winter." *Atmospheric Environment* 24B:57-66.

Lipfert, FW; Wyzga, RE; Baty, JD; Miller, JP. 2006. "Traffic density as a surrogate measure of environmental exposures in studies of air pollution health effects: Long-term mortality in a cohort of US veterans." *Atmos. Environ.* 40(1):154-169.

Mage, DT; Buckley, TJ. 1995. "The Relationship Between Personal Exposures and Ambient Concentrations of Particulate Matter." Paper no. 95-MP18.01. In *Proceedings of the Air & Waste Management Association 88th Annual Meeting, San Antonio, TX,* June.

McDonnell, WF; Nishino-Ishikawa, N; Petersen, FF; Chen, LH; Abbey, DE. 2000. "Relationships of mortality with the fine and coarse fractions of long-term ambient PM_{10} concentrations in nonsmokers." *J. Expo. Anal. Environ. Epidemiol.* 10(5):427-436.

Meng, QY; Turpin, BJ; Polidori, A; Lee, JH; Weisel, C; Morandi, M; Colomoe, S; Stock, T; Winer, A; Zhang, J. 2005. "$PM_{2.5}$ of ambient origin: Estimates and exposure errors relevant to PM epidemiology." *Environ. Sci. Technol.* 39(14):5105-5112.

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

Mudway, IS; Stenfors, N; Duggan, ST; Roxborough, H; Zielinski, H; Marklund, SL; Blomberg, A; Frew, AJ; Sandström, T; Kelly, FJ. 2004. "An *in vitro* and *in vivo* investigation of the effects of diesel exhaust on human airway lining fluid antioxidants." *Arch. Biochem. Biophys.* 423(1):200-212.

National Research Council (NRC). 2009. "Science and Decisions: Advancing Risk Assessment. [The Silver Book.]" Committee on Improving Risk Analysis Approaches Used by the U.S. EPA, National Academies Press (Washington, DC). 421p.

Oberdörster, G. 1996. "Significance of particle parameters in the evaluation of exposure-dose-response relationships of inhaled particles." *Inhalation Toxicology* 8(Suppl):73-89.

Oberdörster, G. 2002. "Toxicokinetics and effects of fibrous and nonfibrous particles." *Inhalation Toxicology* 14:29-56.

Oller AR, Oberdörster G. 2010. "Incorporation of particle size differences between animal studies and human workplace aerosols for deriving exposure limit values." *Regul. Toxicol. Pharmacol.* 57(2-3):181-94.

Ozkaynak, H; Xue, X; Spengler, J; Wallace, L; Pellizzari, E; Jenkins, P. 1996. "Personal exposure to airborne particles and metals: Results from the particle team study in Riverside, California." *J. Expo. Anal. Environ. Epidemiol*. 6(1):57-78.

Pauluhn, J. 2011. "Subchronic inhalation toxicity of iron oxide in rats: Pulmonary toxicity is determined by the particle kinetics typical of poorly soluble particles." *J Appl Toxicol.* Apr 1. doi:10.1002/jat.1668

Peretz, A; Kaufman, JD; Trenga, CA; Allen, J; Carlsten, C; Aulet, MR; Adar, SD; Sullivan, JH. 2008b. "Effects of diesel exhaust inhalation on heart rate variability in human volunteers." *Environ. Res.* 107:178-184.

Peretz, A; Sullivan, JH; Leotta, DF; Trenga, CA; Sands, FN; Allen, J; Carlsten, C; Wilkinson, CW; Gill, EA; Kaufman, JD. 2008a. "Diesel exhaust inhalation elicits acute vasoconstriction in vivo." *Environ. Health Perspect.* 116:937-942.

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; Thun, MJ; Namboodiri, MM; Dockery, DW; Evans, JS; Speizer, FE; Heath, CW. 1995. "Particulate air pollution as a predictor of mortality in a prospective study of U.S. adults." *Am. J. Respir. Crit. Care Med.* 151 :669-674.

Rhomberg, LR; Chandalia, JK; Long, CM; Goodman, JE. 2011. "Measurement error in environmental epidemiology and the shape of exposure-response curves." *Crit. Rev. Toxicol.* 41(8):651-71.

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.

Smith, RL; Spitzner, D; Kim, Y; Fuentes, M. 2000. "Threshold dependence of mortality effects for fine and coarse particles in Phoenix, Arizona." *J. Air Waste Manage. Assoc.* 50(8):1367-1379.

US EPA. 1996. "Air quality criteria for particulate matter, Volume I-III (Final)." Environmental Criteria and Assessment Office (Research Triangle Park, NC), US EPA, Center for Environmental Research Information (Cincinnati, OH), EPA/600/P-95/001aF; EPA/600/P-95/001bF; EPA/600/P-95/001cF. April.

US EPA. 1999. "The Benefits and Costs of the Clean Air Act 1990 to 2010 - EPA Report to Congress." EPA-410-R-99-001. 146p., November.

US EPA. 2002. "Health Assessment Document for Diesel Engine Exhaust." National Center for Environmental Assessment, Office of Research and Development. EPA/600/8-90/057F.

US EPA. 2010. "Quantitative Health Risk Assessment for Particulate Matter (Second external review draft)." Office of Air Quality Planning and Standards, EPA-452/P-10-001. 531p., February.

US EPA. 2011a. "The Benefits and Costs of the Clean Air Act from 1990 to 2020 (Final Report - Rev. A)." Office of Air and Radiation. 238p., April.

US EPA. 2011b. "Policy Assessment for the Review of the Particulate Matter National Ambient Air Quality Standards." Office of Air Quality Planning and Standards, EPA-452/R-11-003, 458p., April.

Utell, MJ; Morrow, PE; Speers, DM; Darling, J; Hyde, RW. 1983. "Airway responses to sulfate and sulfuric acid aerosols in asthmatics." *Am. Rev. Respir. Dis.* 128:444-450.

Valberg, PA; Bruch, J; McCunney, RJ. 2009. "Are rat results from intratracheal instillation of 19 granular dusts a reliable basis for predicting cancer risk?" *Regul Toxicol Pharmacol*. 54(1):72-83.

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.