Technical Comments on the Regulatory Impact Analysis Supporting EPA's Proposed Rule for Utility MACT and Revised NSPS (76 FR 24976)

Anne E. Smith, Ph.D.
NERA Economic Consulting
1255 23rd Street, NW
Washington, DC 20037-1125
August 3, 2011

On May 3, 2011, the US Environmental Protection Agency (EPA) published a Proposed Rule in the Federal Register (76 FR 24976, hereafter called the "Proposed Rule") to set National Emission Standards for Hazardous Air Pollutants (NESHAPs) for hazardous air pollutants (HAPs, also called "air toxics") from coal- and oil-fired steam electric generating units (EGUs) and also to revise the New Source Performance Standards (NSPS) for fossil-fuel-fired electric utility, industrial-commercial-institutional, and small industrial-commercial-institutional steam EGUs. The Proposed Rule is accompanied by and frequently references a Regulatory Impact Analysis (RIA) that EPA released (EPA, 2011a) as the basis for statements about the costs and benefits of the Proposed Rule. I am commenting on the RIA.

Summary of Key Findings in My Comments and My Conclusions

As is further detailed in the rest of my comments, several key findings about the Proposed Rule emerge from a review of EPA's RIA and related EPA documents:

- Although EPA reports that the Proposed Rule will produce annual benefits ranging from $53 billion to $140 billion, these benefits have nothing to do with air toxics at all.

- EPA's estimates of the direct benefits due to reduction of the air toxics that are the specific purpose of this rulemaking range from only $0.0005 billion to $0.006 billion per year[1] – less than .01% of EPA's total benefits estimate – and this is due to reduction of just one of the HAPs, mercury (Hg). EPA concluded it had no basis for estimating benefits from reduction of any of the other EGU HAPs.

- Effectively all of the $53 billion to $140 billion of estimated benefits is due to "co-benefits" from coincidental reductions of fine particulate matter ($PM_{2.5}$), a pollutant that is separately and independently regulated under the Clean Air Act (CAA) as a criteria pollutant.

[1] Stated in a more readable format, the range of benefits estimated for the air toxics is $500,000 per year to $6 million per year. The RIA's summary Table 1-3 incorrectly states the lower bound, and I am reporting the values from RIA Chapter 5 (Table 5-7), and in the Proposed Rule (at 24979).

- The $PM_{2.5}$ co-benefits lack credibility because almost all of that dollar value comes from exposures that are so low that EPA deems them safe and is expected to continue to deem them safe after completing its review of the current $PM_{2.5}$ health standard this year. Further, the reductions in exposure levels are very small, averaging only 0.7 $\mu g/m^3$ in annual average concentrations.[2]

- The $PM_{2.5}$ co-benefits also lack credibility because of a long list of well-documented technical problems with the way EPA chooses to calculate actual health risks from statistical associations that have not been reliably shown to reflect causal relationships. These causality questions are particularly pronounced with respect to individual $PM_{2.5}$ constituents such as sulfate, which is almost the only constituent accounting for the Proposed Rule's co-benefits.

- *Prima facie* evidence of the non-credibility of EPA's co-benefits estimates exists in EPA's baseline estimates of risk in this RIA: deaths that were "due to" ambient $PM_{2.5}$ exposures exceeded 20% in areas of the US in 2005. These co-benefits assumptions also imply that over 40% of deaths were due to $PM_{2.5}$ in parts of the US during the period 1979-1983 when $PM_{2.5}$ concentrations were approximately double those for 2005. These surprisingly high assumptions about baseline risk, which in my opinion stretch the bounds of plausibility, are the result of a single assumption change in 2009 in EPA's RIAs to extrapolate risks below the ambient $PM_{2.5}$ levels that have been studied, to as low as background (*i.e.,* nearly zero).

    - RIAs are not subject to peer review by EPA's Clean Air Scientific Advisory Committee (CASAC) or to a public comment period.

    - EPA has not made this assumption change in any of the risk analyses supporting its current review of the $PM_{2.5}$ health standard, which are subject to CASAC review.

- The $PM_{2.5}$ co-benefits estimates are virtually all tied to attainment of the Proposed Rule's MACT for acid gases, which is the one MACT category in this Proposed Rule for which EPA has not offered any evidence of health risk.

- Given that almost all of the co-benefits are solely attributable to the acid gas MACT portion of the Proposed Rule, there is no cost-benefit case for the remainder of the HAPs control requirements in the rule, whether their estimated co-benefits are included or not.

---

[2] RIA, p. 4-5. To put this in context, the annual average standard (*i.e.,* the level protective of public health with an adequate margin of safety) is 15 $\mu g/m^3$, about 20 times larger. Even the maximum decrease in $PM_{2.5}$ projected under the Proposed Rule is only 1.49 $\mu g/m^3$ (*ibid.*).

In light of the above points, which are further elaborated in the rest of my comments, I conclude that the lower bound of the $PM_{2.5}$ co-benefits should be zero, and that EPA's upper bound $PM_{2.5}$ co-benefits estimate is just not credible. EPA has not even quantified any benefits for the HAPs themselves, other than a tiny benefit from Hg reduction.

More importantly, I conclude that EPA's argument that there is a strong cost-benefit justification for the Proposed Rule is inappropriate because it is based solely on a preponderance of co-benefits from a pollutant that is already regulated, and not an air toxic. Moreover, the estimate is almost entirely derived from changes in very low concentrations that EPA has deemed adequately protect the public health. In the meantime, EPA has not been able to quantify, or even clearly identify, any meaningful amount of direct benefits from the reductions in air toxics that this rule mandates. The maximum ratio of direct benefits to costs for all three MACT groupings is 0.0006-to-1, with a net loss of about $10.9 billion per year. Each individual MACT grouping appears to impose a net benefit-cost loss on the basis of its direct benefits only, and two of those groupings appear to impose net losses even if their share of the upper bound estimates of co-benefits is included in the net benefit calculation.

I am not commenting on the cost analysis in the RIA, but that does not imply that I agree with or endorse the cost and economic impact portions of the RIA.

All of the estimates of costs and benefits cited in my comments are in 2007 dollars, consistent with the RIA.

I. The Proposed Rule's Estimated Benefits Are Not Due to Air Toxics Reductions

EPA reports that the Proposed Rule will produce annual benefits of 6,800 to 17,000 avoided premature deaths and other types of health effects reductions, with an estimated value ranging from $53 billion to $140 billion, but these benefits have nothing to do with air toxics at all. The fact that none of these benefits are due to air toxics reductions is quite clear if one reads the Executive Summary (Chapter 1) of the RIA. However, this fact is completely obscured by misleading rhetoric from groups such as the Natural Resources Defense Council (NRDC) that imply that these are benefits of the air toxics reductions themselves. For example, John Walke of NRDC has testified:

> "EPA's proposed mercury and air toxics standards for power plants that burn coal and oil are projected to save as many as 17,000 American lives every year by 2015. These standards also will prevent up to 11,000 cases of heart attacks, 120,000 cases of asthma attacks, 11,000 cases of acute

> bronchitis among children, 12,000 emergency room and hospital visits and 850,000 lost work days every year."[3]

The numbers NRDC cites in the misleading quote above are not due to mercury or air toxics reductions, but due to a criteria pollutant, $PM_{2.5}$, that EPA already regulates. By careful sentence construction, however, groups like NRDC sway their lay audiences into believing that these health impacts are due to the air toxics themselves, rather than $PM_{2.5}$, thus falsely making the regulation of the air toxics appear to be a public health imperative.

Given that the health impacts cited above are not due to air toxics, it is instructive to ask the question: *What risk reductions has EPA identified for the Proposed Rule's reductions of the air toxics themselves?* EPA's estimates of the benefits due to reduction of air toxics that is the purpose of this rulemaking range from only $0.5 million to $6 million per year, and these estimates are due to reduction of just one of the HAPs being regulated – Hg.[4] These air toxics benefits compare to EPA's $10.9 billion estimate of their cost. The Hg benefits are so low because after exhaustive and comprehensive evaluation of the Hg risks to the most sensitive population – children exposed *in utero* to high methylmercury concentrations – EPA has estimated that the imposition of the Proposed Rule would improve the average IQ of those children by only 0.00209 IQ points.[5] Such a change would not even be measurable in actual IQ testing (the average person's IQ score being 100). The RIA's Table 1-2 which summarizes the physical effects that lie beneath the monetized benefits estimates does not mention this tiny change, but instead provides a meaningless "sum of total lost IQ points" of 510.8 IQ points.[6] Even aggregated in this way, the impact still seems small, given that the comparable sum of total IQ points among all children born each year is about 425 million.[7] It is small even compared to the total IQ points among the 244,000 exposed children, who would have over 24 million IQ points in aggregate.

Although EPA never reports the IQ loss of a child born to a mother who eats recreationally-caught freshwater fish in quantities at the 95th percentile level, it appears that EPA may have also computed this IQ loss because the RIA mentions that EPA assumed 25 gm/day of fish consumption for the 95th percentile consumption level of recreational fishers, while assuming 8 gm/day for that population's average consumption level.[8] The 95th percentile of IQ loss within the sensitive population is thus easily

---

[3] John D. Walke, Natural Resources Defense Council, Testimony at Hearing on "Recent EPA Rulemakings Relating To Boilers, Cement Manufacturing Plants, And Utilities," before the Subcommittee on Energy and Power, Committee on Energy and Commerce, U. S. House of Representatives, April 15, 2011.

[4] RIA, Table 1-3, p. 1-5. However, Table 1-3 incorrectly states the lower bound as $5000, not $500,000; the lower bound I am reporting is in RIA Chapter 5 (Table 5-7), and in the Proposed Rule (at 24979).

[5] RIA, p. 5-2.

[6] RIA, Table 1-2, p. 1-4.

[7] This is calculated by multiplying the number of births in the US each year (about 4.25 million) by the average of 100 IQ points per person.

[8] RIA, p. 5-61.

computed because increased fish consumption affects the estimated maternal Hg intake linearly.[9] Since 25 gm/day is about 3.13 times 8 gm/day, the 95th percentile child's IQ loss would be about 3.13 times .00209, or 0.007 IQ points. Even the 95th percentile IQ loss estimate is smaller than anything that can be detected in IQ testing. (The aggregate total of 510.8 lost IQ points is unaffected, because it is based on the average level regardless of the 95th percentile level.)

As small as the average change per exposed child appears to be, EPA nevertheless assigns projected earnings losses to that change, with the resulting estimate of between $0.5 million and $6 million in the present value of their earning power improvement as a result of the total Hg reduction that would occur under the Proposed Rule.

Even these small Hg benefit estimates are clearly overstated, because EPA assumes that the entire reduction in fish tissue will occur instantaneously with the abatement of EGU emissions, and hence that the IQ benefits will occur in full, by 2016. EPA's RIA acknowledges this is not a sound assumption, saying that its mercury benefits modeling:

> "does not provide for a calculation of the time lag between a reduction in mercury deposition and a reduction in the MeHg concentrations in fish and, as noted earlier, depending on the nature of the watersheds and waterbodies involved, the temporal response time for fish tissue MeHg levels following a change in mercury deposition can range from years to decades."[10]

The footnote EPA attaches to the above quote adds:

> "If a lag in the response of MeHg levels in fish were assumed, the monetized benefits could be significantly lower, depending on the length of the lag and the discount rate used."[11]

This means that any alternative, more realistic assumptions would have produced even lower monetized benefits for Hg. The Proposed Rule, however, is less forthright about these limitations. It relegates mention of this limitation in the Hg benefits estimates to a footnote that itself obscures the potentially significant overstatement of the benefits with the following, more oblique wording:

> "The risk assessment is not designed to track the detailed temporal profile associated with changes in fish tissue MeHg levels following changes in Hg deposition. Rather, we are focusing on estimating risk in the future, assuming that near steady state conditions have been reached (following a simulated change in Hg deposition). Additional detail regarding the temporal profile issue and other related factors (*e.g.*, methylation potential

[9] RIA, equation 5.4, p. 5-60.

[10] RIA, p. 5-25.

[11] RIA, p. 5-25.

> across watersheds) is discussed in Section 1.3 and in Appendix E of the National Scale Mercury Risk Assessment TSD)."[12]

Mercury benefits may be small even with their overstatement, but the RIA was unable to quantify any benefits at all for any of the acid gas, metallic, or organic HAPs reductions. EPA has not even identified any actual health risk associated with current levels of the acid gases. Most of the qualitative hazards that EPA describes for these toxics are for exposures at occupational not ambient levels.[13]

To sum up, estimated benefits for reducing the air toxics that are the purpose of the Proposed Rule ("direct benefits") are between 0.0004% and 0.011% of the total benefits that EPA is attributing to this rule. The RIA states that EPA believes there are substantial unquantified benefits, "including the overall value associated with HAP reductions" and points to the RIA's Table 1-4 for a list of these unquantified HAP reduction benefits.[14] However, RIA Table 1-4 lists only PM health, PM welfare, ozone health, ozone welfare, NO_2 health, NO_x welfare, mercury health, and mercury wildlife effects. Of these, only mercury is an air toxic. The rest of the unquantified benefits listed are co-benefits not related to air toxics exposures. Not one unquantified benefit is listed for acid gases, non-Hg metallic HAPs or organic HAPs. Perhaps the most telling fact of all is that discussion of risks from non-Hg HAPs consumes only 6.5 pages of the 469 pages of the RIA.

Thus, as can also be seen from RIA's Table 1-3 on (pp. 1-4 to 1-5 of the RIA), effectively all of the $53 billion to $140 billion of estimated benefits are "co-benefits" from reductions of pollutants *that are not air toxics at all* but which EPA estimates also will be reduced in the course of efforts to control the air toxics to meet their proposed new standards. Of this total, fully $52.4 billion to $139.4 billion is due to co-benefits from a single ambient pollutant – $PM_{2.5}$ – which is already the subject of health-protective regulation by EPA. (The remaining $0.6 billion of co-benefits is an estimate of the social benefit of reduced greenhouse gases ("carbon"), which comes from reduced coal-fired generation under the Proposed Rule due to projected coal plant retirements as a result of the costs of compliance.)

As I will explain in Section III, the $PM_{2.5}$ co-benefits that completely dominate the RIA's benefits estimates are dubious because they are based on changes in $PM_{2.5}$ exposures that are, even before being lowered, at levels that EPA deems safe. I then explain three additional broad sets of reasons that they are not credible in Section IV. However, before reading further, readers unfamiliar with the literature on $PM_{2.5}$ health risks should be aware that the estimates of $PM_{2.5}$-attributed deaths (such as the 6,800 to 17,000 that EPA is attributing to the Proposed Rule) are based entirely on statistical associations between total mortality rates in various locations of the US and their respective monitored, region-

[12] Proposed Rule, footnote 99, p. 25009.

[13] See, *e.g.*, Proposed Rule, pp. 25003-25005.

[14] RIA, p. 1-8.

wide ambient $PM_{2.5}$ concentrations. As I will demonstrate in Section IV.1, EPA's estimate of 6,800 to 17,000 $PM_{2.5}$-related premature deaths avoided in 2016 as a result of the Proposed Rule is based on an assumption that 130,000 to 320,000 deaths, respectively, of 2005's US deaths were hastened by breathing ambient $PM_{2.5}$.[15] And yet, EPA identifies not a single death during 2005 that was attributed, even in part, to exposure to ambient $PM_{2.5}$. If $PM_{2.5}$ is indeed having this estimated effect on the public health, there is no evidence indicating when or where these events occurred, or who was affected. Rather, these mortality estimates are merely inferences drawn after making a host of assumptions about how to convert a statistical association into a concentration-response function. No one really even knows what types of deaths might be implicated. A common belief among researchers is that the deaths are primarily cardiovascular in nature, but this is far from an established fact: everything from cardiovascular causes to diabetes to lung cancer has been mentioned as having such an association in one paper or another. There is no clinical evidence to inform these inferences either, despite at least 15 years of efforts by researchers to find a clear physiological mechanism to explain and lend credibility to these estimates based solely on statistical correlations.

As the hackneyed phrase taught to all introductory statistics students goes: "Correlation does not imply causation." All of the risk estimates that the RIA attributes to $PM_{2.5}$ are based on a *presumption* that the associations in the epidemiological literature are causal in nature. This presumption remains under debate, as I will explain in Section IV.2.

I have emphasized the mortality evidence just now, and continue to do so in the rest of my comments because almost all of the co-benefits are for mortality. However, the use of only statistical associations to infer and then project changes in health outcomes applies to all of the morbidity (*i.e.*, non-fatal) health benefits estimates as well.

II. No Cost-Benefit Case Exists for Any of the HAPs Groupings Regulated by the Proposed Rule

Given that EPA relies entirely on $PM_{2.5}$ co-benefits to create the appearance that the Proposed Rule provides much greater benefits than costs when reducing air toxics, it is useful to examine whether these co-benefits are associated with any specific part of the new regulations that would be imposed under the Proposed Rule. A key feature of the NESHAPs portion of the Proposed Rule is that it sets a standard of Maximum Achievable Control Technology (MACT) for a variety of HAPs.[16] The Proposed Rule sets MACTs in three generic groupings: for Hg, for the entire group of non-mercury metallic HAPs

[15] Fann *et al.* (2011), Table 1, p. 8.

[16] A NESHAP, including for EGU HAPs, does not necessarily have to be based on MACT. In fact, the Proposed Rule would regulate organic HAPs (*e.g.*, formaldehyde) with a work practice standard rather than a MACT-based standard (Proposed Rule, p. 25027). Nevertheless, the Proposed Rule is frequently referred to informally as the "EGU MACT Rule," and the bulk of my comments focuses on the MACT standards in the Proposed Rule.

(using particulate matter emissions as a surrogate),[17] and for the entire group of acid gases (using hydrogen chloride (HCl) as a surrogate).[18] EPA grouped the HAPs in this manner because the Agency found that the HAPs within each group can be most effectively controlled by a single type of technology that differs for each group. For example control of non-Hg metal HAPs occurs primarily through particulate control devices, while control of acid gases is generally achieved using some form of flue gas desulfurization technology. Hg is more complex because several types of technology may be effective, but the most cost-effective on a stand-alone basis is activated carbon injection (ACI), which is uniquely targeted to capturing Hg.

Thus, EPA has performed a separate MACT analysis for each of these three groups of HAPs, and a work practice standard for non-Hg organic HAPs. Estimates of benefits and benefit-cost comparisons therefore must vary by group, but EPA has not provided such group-specific cost and benefit information. EPA has provided an approximate breakout of costs for individual pollutant controls.[19] However, it has not provided a comparable breakout for its benefits, which is also needed to obtain insights about the costs and benefits broken out to the three MACT groupings. I have therefore made my own approximation of the breakout of benefits, which is presented in Table 1 below. I provide the details of how I made the co-benefits disaggregation in Appendix A of this document.

Table 1. Approximate Attribution of Costs, Benefits, and Co-Benefits by Element of Proposed Rule
(Costs from Proposed Rule Table 25; See Appendix A for Details of Benefits Disaggregation)

	(a) Benefits from air toxics reduction (billions/yr)	(b) Co-benefits from non-toxics (billions/yr)	(c) Costs (billions/yr)	(d) Net Benefits <u>without</u> co-benefits (billions/yr)	(e) Net Benefits <u>including</u> co-benefits (billions/yr)
Mercury MACT	< $0.1	$0.6 to $1.5	$2.3	- $2.3	- $1.7 to -$0.8
Acid Gases MACT	$0	51.7 to 136.9	$5.4	- $5.4	$46.2 to $131.5
Non-Hg Metals MACT	$0	0.7 to 1.6	$3.2	- $3.2	- $2.5 to -$1.6
Organic HAPs Standard	$0	$0	>$0(*)	<$0	<$0
Total	< $0.1	$53 to $140	$10.9	- $10.9	$42 to 129.1

(*) EPA has not provided an overall estimate of the cost of the organic HAPs work practices requirement. However Table 14 of the Proposed Rule (p. 25052), EPA indicates that "tune ups" will cost about $3000 each annually, plus about $17,000 in capital costs (presumably per plant), clearly indicating that EPA does not view this provision as costless.

[17] The metallic HAPs of greatest concern as risk drivers in this Proposed Rule are chromium VI (Cr^{+6}), arsenic (As), and nickel (Ni) (Strum *et al.*, 2011, Table 5, p. 15).

[18] The acid gas of greatest concern as a risk driver in this Proposed Rule is HCl (Strum *et al.*, 2011, Table 5, p. 15).

[19] Proposed Rule, Table 25, p. 25075.

Column d of Table 1 reveals that all of the portions of the Proposed Rule have negative net benefits (*i.e.* their costs are greater than their benefits) if co-benefits are not included. That net negative benefit is several billions of dollars per year for each of the three MACT groups. However, it is also very interesting that even if co-benefits are included (see column e), only the acid gases MACT group has a positive net benefit, while the MACTs for Hg and for the non-Hg metallic HAPs still have negative net benefits. As for the acid gases, if co-benefits *are* included, this group of HAPs is in the remarkable position of being viewed as passing a cost-benefit test by a vast margin, despite billions of dollars of cost and zero dollars of identified direct benefits from the acid gas reductions. This bizarre result merits further discussion.

The huge net benefits that are estimated for the acid gases MACT group occur because almost all of the $PM_{2.5}$ co-benefits that EPA has estimated are due to reductions in the *sulfate* component of ambient $PM_{2.5}$. This, in turn, is almost entirely attributable to the requirement to reduce acid gases through installation of some form of flue gas desulfurization technology, which also reduces SO_2.[20] In contrast, EPA reports that reductions of ambient $PM_{2.5}$ concentrations resulting from incremental reductions of primary $PM_{2.5}$ emissions reductions due to the Proposed Rule are "very small" compared to the sulfate reductions.[21]

Thus, the RIA's inclusion of co-benefits predominantly helps build a cost-benefit case for the acid gases MACT category, which is notably the one MACT grouping for which EPA has not offered any evidence of direct health effects. For example, none of the acid gases is listed as carcinogenic. Further, in its inhalation risk analyses, EPA has estimated hazard quotients (HQs) for HAPs that pose non-cancer health risks from chronic exposure. EPA states that if an HQ is 1.0, estimated exposures are at a level "that is likely to be without an appreciable risk of deleterious effects during a lifetime,"[22] but above that point, EPA considers the margin of safety against toxic effects to be too uncertain to be acceptable. EPA reports that the HQ for HCl never exceeded 0.05 in any of its inhalation risk estimates,[23] meaning that for EGUs, the predominant HAP in the

---

[20] The SO_2 reductions must be beyond what existing standards, including the $PM_{2.5}$ NAAQS and the SO_2 NAAQS, will require in order to be appropriate to consider as co-benefits rather than merely double-counted. Whether that is the case or not is not clearly demonstrated, as I will discuss in Section IV.3.

[21] RIA, p. 6-11. EPA also reports that nitrate $PM_{2.5}$ actually increases, but has not included this negative co-benefit in its co-benefits calculation. The effect is also reported to be "small" relative to sulfate reductions, but it may be larger than the "very small" share of co-benefits due to direct $PM_{2.5}$ emissions reductions.

[22] Strum *et al.* (2011), p. 13.

[23] Proposed Rule, footnote 170, p. 25051. Although EPA notes that other acid gases (Cl_2, HF and HCN) were not included in the risk calculations "because of uncertainties in their emission rates," it is hardly likely that any of these other gases would involve a HQ so much closer to 1.0 than HCl, given that their total EGU emissions are less than 15% of total EGU HCl emissions (see Table 4 of Proposed Rule, p. 25005).

acid gas MACT group has a maximum risk that is 95% below a level that EPA deems protective of health with a safety factor included.

Neither does EPA present any evidence that further controls of acid gases would benefit ecosystems; the Agency merely refers to such impacts as a possibility:

> "In areas where the deposition of acids derived from emissions of sulfur and NO_x are causing aquatic and/or terrestrial acidification, with accompanying ecological impacts, the deposition of hydrochloric acid could exacerbate these impacts. Recent research has suggested that deposition of airborne HCl has had a greater impact on ecosystem acidification than previously thought, although direct quantification of these impacts remains an uncertain process."[24]

Thus, as Table 1 shows, EPA has not been able to quantify any direct benefits from controlling the acid gas HAPs, and this is because it could not find any evidence of current acute or chronic health risks from EGU emissions of these gases. Section 112(d)(4) of the CAA gives EPA discretion to consider setting a "health-based" standard for a HAP that has an HQ below 1.0. A health-based standard can be less stringent (and less costly) than MACT, provided that it protects health with an ample margin of safety (for example, by ensuring HQs will be lower than 1.0). As the Proposed Rule notes, EPA has applied health-based standards for HCl under Section 112(d)(4) in other NESHAP rulemakings.[25]

III. Air Toxics Regulation Should Not Be Justified on the Basis of Co-Benefits from Coincidental Reductions of a Separate, Already-Regulated Pollutant

The fact that the cost-benefit case for the acid gas MACT is based solely on co-benefits, and not on any risks from those air toxics themselves raises an important question for EPA's entire cost-benefit case:

> *Is it appropriate to claim positive net benefits from controlling acid gases when there are no identified health risks or other evidence of benefits from reducing the acid gases themselves and all of the reason for positive net benefits are co-benefits from another pollutant whose health risks are already separately and independently regulated by EPA?*

I do not consider this an appropriate application of cost-benefit analysis. It is illogical to control a pollutant for which there are no known direct benefits solely on the basis of the co-benefits from controlling it. This is particularly illogical when the pollutant that accounts for those co-benefits is already regulated directly, as is the case with $PM_{2.5}$. Also, relying on $PM_{2.5}$ co-benefits to justify establishing EGU air toxics regulations could

[24] Proposed Rule, p. 25050, footnote omitted, emphasis added.

[25] Proposed Rule, p. 25050.

effectively put the entire burden of lowering $PM_{2.5}$ levels to meet a future NAAQS on a single industry and pre-empt the SIP process. That would likely impose higher-cost compliance than the SIP process is supposed to provide and also place an unfair burden on one sector, the electric sector.

$PM_{2.5}$ risks are already required to be controlled to levels that EPA deems protective of the public health with an adequate margin of safety under the $PM_{2.5}$ NAAQS.[26] Yet, effectively all of the $53 billion to $140 billion in $PM_{2.5}$ co-benefits are due to reductions in exposures to $PM_{2.5}$ that are already below the annual NAAQS standard of 15 $\mu g/m^3$. This fact can be inferred from Figure 6-15 of EPA's RIA (copied as Figure 1 on the next page) in the following way. The blue S-shaped curve in Figure 1 indicates on the vertical axis the percent of the RIA's $PM_{2.5}$ co-benefits estimate that is attributable to baseline $PM_{2.5}$ exposures at or below the $PM_{2.5}$ concentration on the horizontal axis. This is known as a "cumulative distribution." The point on the horizontal axis where the S-shaped curve just reaches 100% indicates the level of baseline $PM_{2.5}$ at or below which *all* (*i.e.*, "100%") of the estimated $PM_{2.5}$ co-benefits are attributable. I have added a vertical dotted red line to Figure 1 at the level of the current annual NAAQS (*i.e.*, at 15 $\mu g/m^3$ on the horizontal axis). As one can see, the vertical reading on the blue S-shaped curve is essentially 100% at 15 $\mu g/m^3$, which means that effectively all of EPA's estimated $PM_{2.5}$ co-benefits are based on reductions in baseline $PM_{2.5}$ exposures that are already below the current health-protective air quality standard for long-term $PM_{2.5}$ exposures, and which therefore EPA has deemed do not impose unacceptable public health risks.

EPA is presently considering whether to tighten the $PM_{2.5}$ NAAQS, with a Proposed Rule expected later in 2011. EPA is considering a range of possible alternative annual standards that extends as low as 11 $\mu g/m^3$. If EPA revises the NAAQS to the lowest of those levels, the RIA's Figure 6-15 (Figure 1 above) tells us that 20% of the co-benefits being attributed to the acid gases MACT (*i.e.*, those that occur in locations where pre-rule $PM_{2.5}$ is above 11 $\mu g/m^3$) are going to occur anyway, as a result of NAAQS compliance.[27] They therefore would be inappropriate to count as co-benefits of the Proposed Rule for air toxics – they should be counted as the direct benefits of the new $PM_{2.5}$ standard.

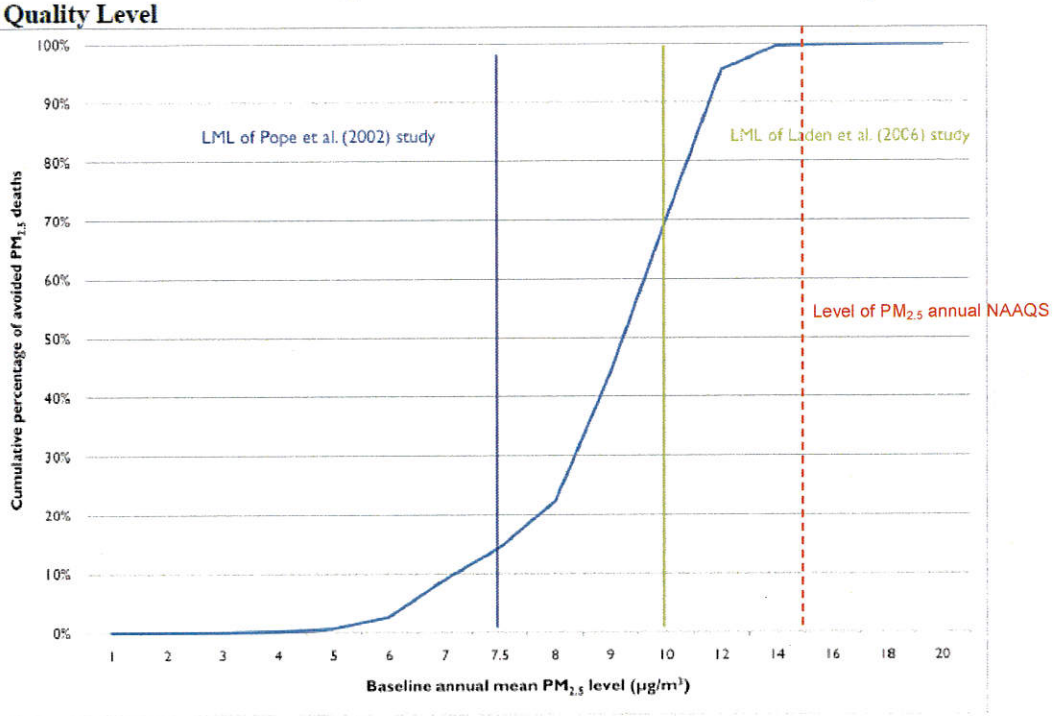
---

[26] The current $PM_{2.5}$ NAAQS is a limit of 15 $\mu g/m^3$ for annual average $PM_{2.5}$ and a limit of 35 $\mu g/m^3$ for the 3-year average of the 98th percentile 24-hour average concentration of $PM_{2.5}$. It was set in 2006, at which time the 24-hour average limit was reduced from 65 $\mu g/m^3$ to 35 $\mu g/m^3$, and the benefits for that tightening were calculated in that rule's RIA (*i.e.*, EPA, 2006).

[27] Some might argue that these $PM_{2.5}$ benefits will appear sooner because the Proposed MACT Rule will be fully implemented by 2016, while full implementation of a tightened $PM_{2.5}$ NAAQS will be several years later. However, that difference is only temporary, and many have argued that the accelerated time frame for implementation of the EGU MACT rule will be far more disruptive than EPA's cost analysis indicates due to its exceedingly rapid implementation. Thus, making a point that these could be considered valid *temporary* co-benefits for the years 2016 through perhaps 2020 only raises the question of whether that accelerated time frame is reasonable and justifiable.

Figure 1. Copy of Figure 6-15 from the RIA
(The dotted red vertical line has been added to identify the level of the current annual $PM_{2.5}$ NAAQS)

Figure 6-15. Cumulative Percentage of Total PM-Related Mortalities Avoided by Baseline Air Quality Level



Of the total PM-related deaths avoided:
86% occur among population exposed to PM levels at or above the LML of the **Pope et al.** study.
30% occur among population exposed to PM levels at or above the LML of the **Laden et al.** study.

The more important point, however, is that if it were viewed as credible that such large effects exist below the level of the standard, the appropriate remedy would be to tighten the $PM_{2.5}$ standard, and not to regulate something else altogether in order to obtain those benefits through "coincidence." However, even EPA and its advisors on CASAC do not appear to have enough confidence in the scientific evidence to argue that a $PM_{2.5}$ standard at or below that level is requisite to protect the public health.[28] Clearly, the evidence from EPA's own actions and stated intentions is that the remaining 80% of the co-benefits are not credible enough to justify a tighter $PM_{2.5}$ NAAQS. How then could these estimates possibly be credible enough to justify spending billions of dollars per year on controls on acid gases when EPA cannot identify any credible evidence of current public health risk from the acid gases themselves?

There are, however, numerous other reasons the $PM_{2.5}$ co-benefits in the Proposed Rule should be given little credence, which I describe in the next section.

[28] EPA (2011b), p. 2-106.

IV. Three Categories of Reasons Cast Doubt on the Credibility of the $PM_{2.5}$ Co-Benefits Estimates

If the reader is inclined to accept a preponderance of co-benefits from an already-regulated pollutant to justify a regulation for a different and unrelated pollutant, he/she should then also become informed about the many, more technical reasons to doubt the credibility of the $PM_{2.5}$ co-benefits estimates in this RIA. That is the topic of this section, which covers three general areas of concern:

- Implausibly high baseline risk estimates for $PM_{2.5}$
- Technical gaps and limitations that remain in the epidemiological evidence
- Double-counting of $PM_{2.5}$ co-benefits across multiple RIAs

1. EPA's Baseline Risk Levels Provide *Prima Facie* Evidence of Non-Credible $PM_{2.5}$ Risk Estimation Methods

My points thus far regarding the lack of credibility of the co-benefits estimates in the air toxics Proposed Rule have been based on evidence from EPA and CASAC's own choices and behaviors with respect to $PM_{2.5}$ regulation that reveals their own understanding of the weakness in the evidence supporting these hypothetical calculations of co-benefits. However, there is other information in the RIA that sheds light on the *prima facie* problems of credibility with EPA's co-benefits estimates. To wit, data buried in the RIA (and supported by other related papers and data) show that EPA is starting from an assumption for the upper bound co-benefits estimate that implies that as recently as 2005, the act of breathing ambient levels of fine particles on a daily basis was a contributing factor in over 20% of *all* deaths in parts of the US. This is a startlingly high percentage, particularly given that by 2005 air quality was much better than in the immediately preceding decades. For example, as I will explain below, EPA's upper bound estimate is also implying that as recently as 1979-1983 over 40% of deaths were due to breathing the then-current ambient levels of $PM_{2.5}$ in some US urban areas, and that the *nationwide average* of all deaths due to ambient $PM_{2.5}$ was about 25% (one in four).

The evidence in the RIA that starts to bring these implausible assumptions to light appears in Figure C-2 from p. C-11 of Appendix C of the RIA. I have copied it into my comments as Figure 2 on the next page.

Figure 2. Copy of Figure C-2 from the RIA.

Figure C-2. Distribution of $PM_{2.5}$ Mortality Risk in 2005

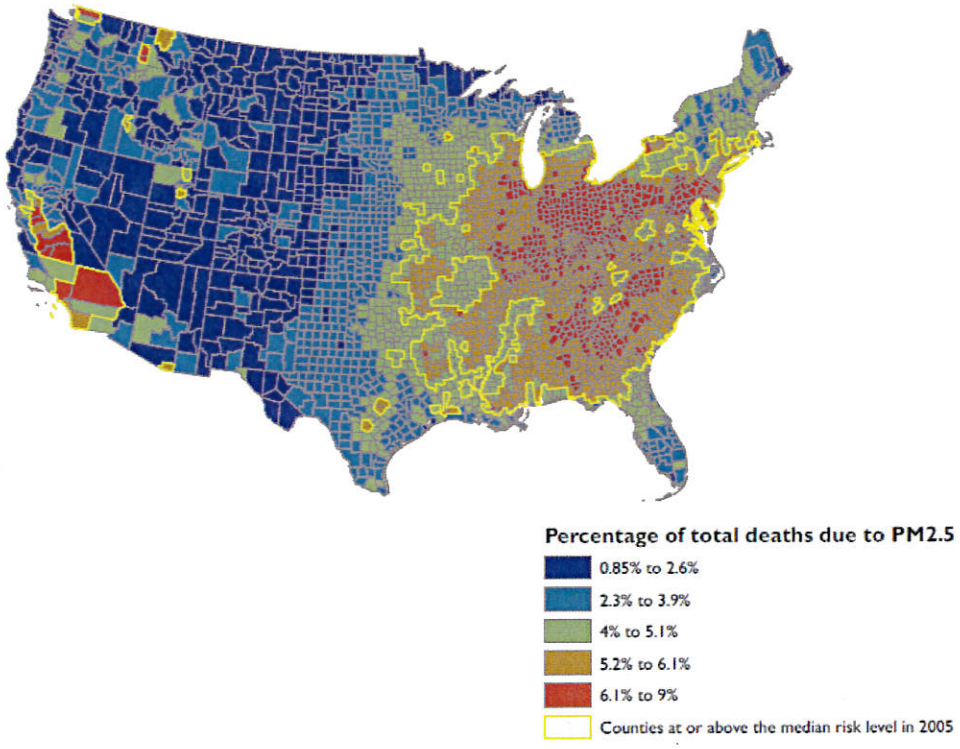


Figure 2 (the RIA's Figure C-2) shows that for many counties in the Eastern US, EPA finds that 6.1% to 9% of all deaths are "due to $PM_{2.5}$," while for most of the rest of the East, 5.2% to 6.1% of all deaths are due to $PM_{2.5}$. This particular figure does not indicate what I stated above, that up to 20% of deaths are due to $PM_{2.5}$; to realize that, one must "read the fine print" that accompanies this figure in the RIA. When one does, it becomes apparent that Figure C-2 of the RIA is not based on the same relative risk assumptions that EPA has used to generate its high-end mortality estimates. This is important because EPA emphasizes only the high-end estimates in its public statements.[29] Specifically, the RIA says that the calculations used to produce its Figure C-2 used a relative risk of 1.06

[29] I already quoted an NRDC summary earlier in my comments, which mentioned only the upper bound of deaths (*i.e.*, 17,000). EPA does the same. In its announcement that it would provide a 30-day extension on the comment period for the Proposed Rule, the Administrator also only mentioned its upper bound mortality risk estimate: "When these new standards are finalized, they will assist in preventing 11,000 heart attacks, 17,000 premature deaths, 120,000 cases of childhood asthma symptoms and approximately 11,000 fewer cases of acute bronchitis among children each year. Hospital visits will be reduced and nearly 850,000 fewer days of work will be missed due to illness." (EPA Air News Release (HQ), "EPA Extends Public Comment on Mercury and Air Toxics Standards," June 21, 2011).

per 10 μg/m^3.[30] Only the lower bound of the RIA's stated benefits range is calculated using a relative risk of 1.06 (*i.e.,* the estimate based on Pope *et al.*, 2002[31]) while EPA's upper bound estimate is calculated using a relative risk of 1.16 (*i.e.* it is based on Laden *et al.*, 2006[32]).[33] If the upper bound estimate's relative risk were used, the areas in Figure 2 that are shaded red would be labeled as experiencing between about 16% and 22% of all US deaths due to $PM_{2.5}$ in 2005![34] The RIA does not show this or mention it.

Another thing that EPA does not show or mention in the RIA is that its upper bound estimate implies that 13% of all deaths throughout the US in 2005 were "due to $PM_{2.5}$." This fact can be found in a forthcoming publication by EPA authors that the RIA's Appendix C mentions as a source document for its calculations behind its Figure C-2 (*i.e.*, Fann *et al.*, 2011). Fann *et al.* (2011) provides the number of total deaths estimated due to $PM_{2.5}$ using the Laden *et al.* study that is the basis for the upper bound co-benefits in

---

[30] The RIA states at p. C-5: "We substitute risk estimates drawn from the Krewski *et al.* (2009) extended analysis of the ACS cohort. In particular, we applied the all-cause mortality risk estimate random effects Cox model that controls for 44 individual and 7 ecological covariates, using average exposure levels for 1999-2000 over 116 U.S. cities (Krewski *et al.* 2009) (RR=1.06, 95% confidence intervals 1.04—1.08 per 10 μg/m^3 increase in $PM_{2.5}$)."

[31] See relative risk of 1.06 per 10 μg/m^3 increase in $PM_{2.5}$ for "all-cause mortality", average over all years reported in Pope *et al.* (2002), Table 2, p. 1136.

[32] See relative risk of 1.16 per 10 μg/m^3 increase in $PM_{2.5}$ for "total mortality", average over entire follow-up period, reported in Laden *et al.* (2006), Table 3, p. 670.

[33] The ratio of the difference in those two relative risks (*i.e.*, .16/.06) accounts for the difference in the upper and lower bound $PM_{2.5}$ mortality co-benefits (*i.e.,* 17,000 deaths/6,800 deaths). To be more technically precise, the "beta coefficients" used in the $PM_{2.5}$ mortality risk calculations are not precisely the same as the stated "relative risks for a 10 μg/m^3 change in $PM_{2.5}$." My Appendix B provides a more thorough explanation of these points, but the beta coefficient for the 1.16 relative risk is approximately 0.015, while it is approximately 0.006 for the 1.06 lower bound. The ratio of the betas (.015/.006) is 2.5, which is almost exactly the value of the ratio of the computed respective deaths (17000/6800).

[34] This is computed by first back-calculating the $PM_{2.5}$ concentration associated with 6.1% and 9% mortality when using the Krewski *et al.* beta coefficient that EPA says it used for Figure 2 (*i.e.,* .00554, which can be found in Table C-1 of EPA (2010b) at p. C-3). For example, 9% of deaths implies an increase over baseline risk of 9.89% (*i.e.*, defining R as the *increase* in risk due to PM, R/1+R equals the fraction of deaths *due to* PM, .09, which means R=.0989). This then implies that exp(PM*.00554)=1.0989, where PM is the increase in $PM_{2.5}$ concentration that produces an increase in baseline risk of 9.89% and the risk coefficient is .00554 from the Krewski *et al.* result that EPA says it used to produce RIA Figure C-2. Solving this equation gives an increment of $PM_{2.5}$ of 17.02 μg/m^3 for the 9% case. To this $PM_{2.5}$ increment, one can now apply the risk coefficient from Laden *et al.*, which is about .015, to calculate the equivalent increase in baseline risk for the upper bound co-benefits estimate: R^{UB}=exp(17.02*.015)-1. Solving this implies an upper bound increase in risk of 0.29, which in turn implies that the percent of total mortality *due to* that $PM_{2.5}$ increment when using the upper bound relative risk is .29/1.29, or 22%. By going through the same steps, but starting with 6.1% instead of 9%, one finds that the upper bound risk estimate's equivalent for 6.1% is 15.7%. Given potential rounding errors in these calculations, I have rounded my estimates down to come up with the range 15% to 22% as being the upper-bound's equivalent to 6.1% to 9%. See my Appendix B for more explanation of these risk calculations.

the RIA: 320,000 deaths in 2005.[35] Given that there were about 2.4 million deaths in the US in 2005, they are reporting that 13% of all 2005 US deaths were due to $PM_{2.5}$.

Consistency between Fann *et al.*'s result and the upper bound estimates in the RIA can be confirmed from the RIA Figure C-2 by again converting results based on the relative risk of 1.06 from Krewski *et al.* to the upper-bound estimates' relative risk of 1.16. Note in Figure C-2 of the RIA (Figure 2 above), which is based on risk calculations using the relative risk from Krewski *et al.*, that the median US risk level in 2005 due to $PM_{2.5}$ was somewhere between 4% and 5.1% of all deaths; that is, the yellow line that delineates the counties that are *at* the national median risk level runs through green-colored counties. This in turn implies that EPA's national median estimate of percent of deaths due to $PM_{2.5}$ for its upper bound estimate would be between about 10% and 13%,[36] which is consistent with the national average being about 13%, as reported in Fann *et al.* (2011).

However, what is most remarkable is that the geographical extent of counties in the red and tan zones in Figure 2 tells us that EPA's upper bound $PM_{2.5}$ risk calculations are implying that the percent of all 2005 deaths due to $PM_{2.5}$ exceeded 13% throughout almost all counties of the Eastern US, rising to 22% in some locations. Fann *et al.* do not appear to find this result implausible, nor do they appear to have contemplated that their analysis also implies that about 25% of *all deaths nationwide* were due to breathing ambient $PM_{2.5}$ as recently as 1979-1983, with the level being over 40% in some cities.[37] EPA's emissions trend data indicate that the major precursor to ambient $PM_{2.5}$ (*i.e.*, SO_2) was substantially higher in the decade prior to 1979-1983,[38] which suggests that EPA's upper bound estimate of the percentage of deaths due to $PM_{2.5}$ for 1970 might be much higher still, if EPA were to calculate it.

These percentages may seem quite astounding to some readers, and some people (including myself), may conclude from these very high baseline percentages that the mortality estimates of 17,000 due to the air toxics rule that EPA derives from these baseline risk estimates are not really credible. Interestingly, EPA does not report its underlying assumptions on the percent of total 2005 deaths due to $PM_{2.5}$ – not even for its lower bound estimate – anywhere except in the RIA's Appendix C. Further, EPA obscures the fact that those results in Appendix C are only comparable to its lower bound

---

[35] Fann *et al.* (2011), Table 1, p. 8.

[36] To derive this, I applied the same calculation described in footnote 34 to convert the values of 4% and 5.1% based on the Krewski *et al.* risk coefficient of .00554 to the equivalent value based on the Laden *et al.* risk coefficient of about .015.

[37] I base these calculations on the Laden *et al.* beta coefficient of about 0.015, and the annual mean $PM_{2.5}$ data for 63 US cities documented in Krewski *et al.* (2000), in which the average concentration across all the cities was 20 $\mu g/m^3$, and the highest value was 38 $\mu g/m^3$ (Krewski *et al.*, 2000, Part II, Table 30, p. 172). The 95th percentile was 29 $\mu g/m^3$, for which EPA's upper bound calculation predicts that about 35% of all deaths would be due to $PM_{2.5}$. My Appendix B provides more explanation on how such risk computations are done.

[38] See the SO_2 data from 1970-2008 available at http://www.epa.gov/ttn/chief/trends/index.html.

estimate by citing its use of a 1.06 relative risk to a report other than Pope *et al.* (2002), and not offering readers the helpful information that this alternative study (*i.e.,* Krewski *et al.*, 2009) happens to have approximately the same relative risk estimate as Pope *et al.* Understanding that Appendix C's figures reflect only the lower-bound baseline risk levels is therefore left to the highly persistent and diligent reader of the RIA, and the shock factor regarding the implications of the upper bound estimates is preserved for those who actually know the relative risks in the original papers on which EPA's upper and lower bound mortality risks are based. Given that EPA usually only mentions the upper bound estimates of deaths (*i.e.,* 17,000 deaths) in its public statements regarding the benefits of this rule, it is also quite inappropriate that Appendix C does not actually present comparable information regarding the underlying baseline incidences for its upper bound benefits estimates. It certainly could have done so, as it is obvious from Fann *et al.* (2011) that EPA has already performed those calculations.

Some might respond that EPA is not doing anything differently in its $PM_{2.5}$ risk analysis methods than it has been doing since it started such risk analyses in 1996. This would not be true, however. Starting in 2009, EPA decided its RIAs would quantify risks for exposures to $PM_{2.5}$ at levels below the lowest measured levels (LMLs) of $PM_{2.5}$ available for the epidemiological studies.[39] Extrapolating risks below the LML had the effect of inflating the baseline with $PM_{2.5}$-related deaths risks by now including small risks for the vast majority of US exposures that fall below the lowest measured $PM_{2.5}$ level in Laden *et al.* of 10 $\mu g/m^3$ (see Figure 1 above). In my Appendix B, I explain why this is an inappropriate action to take, given the purely statistical basis for the risk estimates. With this single inappropriate change in its calculation methods, EPA went from assuming (when relying on the higher-end Laden-based risk calculation) that among people exposed to $PM_{2.5}$ of, say, 12 $\mu g/m^3$, about 3% of all their deaths are due to $PM_{2.5}$ to now assuming that almost 13% of all those same deaths are due to $PM_{2.5}$.[40] This inflationary effect can be observed just by comparing EPA's baseline 2005 risk estimates in its 2010 $PM_{2.5}$ *Quantitative Health Risk Assessment for* $PM_{2.5}$ – which does *not* extrapolate below

[39] USEPA (2010a), p. S3-3. EPA provides no sound basis for this change. At RIA p. 6-86, EPA attributes this change to a statement in the EPA Integrated Science Assessment for Particulate Matter (EPA, 2009) that the statistical analyses that comprise the epidemiological literature consistently find that no-threshold models provide the best fits for the data to which they are applied. This is not a new finding that warrants a change in EPA risk calculation methods. It dates back through the previous NAAQS-related reviews of the $PM_{2.5}$ literature, and during all that time CASAC never suggested to EPA that this means that there *is* no threshold or other non-linearity at the very low $PM_{2.5}$ levels below each study's LML. There was no sudden change in information in EPA (2009) that warranted EPA's decision to start counting benefits not only below the LMLs, but all the way to background concentrations (which are near zero). Notably, EPA does not extrapolate below the LML in its final Quantitative Health Risk Assessment for $PM_{2.5}$ (EPA, 2010b) which is a document subject to CASAC review, unlike RIAs.

[40] That is, the baseline relative risk of death attributable to $PM_{2.5}$, which used to be computed as exp[($PM_{2.5}$-LML)*.015], became exp[($PM_{2.5}$-PRB)*.015], where PRB stands for policy relevant background. Inserting an LML of 10, a PRB of 3, and a $PM_{2.5}$ level of 12 produces 1.04 and 1.15, respectively, which when restated as a percent of total deaths due to the $PM_{2.5}$ is 3% and 13% respectively. See my Appendix B for the basis for these formulas.

the LML – to those in this RIA which *does* extrapolate below the LML. In the former, EPA estimates 88,000 deaths were due to $PM_{2.5}$ in 2005 using Laden *et al.*,[41] while in the current RIA, EPA estimates 320,000 deaths for the same year, the same estimated air quality, and using the same Laden *et al.* study.[42] The former is 3% of total annual US deaths of 2.4 million and the latter is 13% of 2.4 million annual US deaths. Notably, EPA is now using both of these contradictory estimates of baseline $PM_{2.5}$-related deaths simultaneously in different regulatory proceedings – EPA is using the smaller number of baseline deaths in its CASAC-reviewed risk analyses for the $PM_{2.5}$ NAAQS review, and it is using the larger number of baseline deaths in its RIAs that are generating the large co-benefits for non-$PM_{2.5}$ regulations such as the Proposed Rule for air toxics.

The decision to count all risks even below the LML created a significant inflation in benefits, and did so by adding in benefits of the least credible sort because most of that increase is due to benefits estimates below – often far below – the LML.[43] Thus, overnight in 2009, in the course of preparing RIAs that are not subject to public peer review, EPA dramatically escalated its estimates of benefits for all of its RIAs. This had the most profound impact on its estimates of benefits in the vast swath of the US that has $PM_{2.5}$ concentrations below 10 $\mu g/m^3$: small changes in modeled $PM_{2.5}$ in these areas used to contribute nothing to the total estimated benefits of a regulation, but they now contribute fully 70% to the upper bound benefits estimate (see Figure 1). EPA accomplished this enormous benefits inflation without changing the epidemiological studies it relies on, but by altering a much more obscure assumption in its risk analysis calculations.

One associated and interesting effect of this benefits inflation, however, is the degree to which it causes the baseline percentage of all deaths due to $PM_{2.5}$ to start to strain the bounds of credibility. When EPA's upper bound risk calculation indicates that up to 20% of deaths in some parts of the country were due to $PM_{2.5}$ in 2005, and that up to 40% were due to $PM_{2.5}$ in about 1980, this may stretch the limits of credibility of some readers. It is perhaps unsurprising therefore that EPA is no longer reporting these baseline risk levels in its RIAs.

The simple reason why these new baseline risks are so large – implausibly large in my view – is that EPA assumes in its risk analysis calculations that there is no tapering off of relative risk as $PM_{2.5}$ exposure approaches zero. For years there has been a debate about whether the concentration-response relationship can truly be linear down to zero, but this debate has been focused on questions of statistical power and on basic principles of toxicology. The implication of the linear-to-zero/no-threshold assumption has never been

[41] EPA (2010b), p. G-2.

[42] Fann *et al.* (2011), Table 1, p. 8.

[43] Not only are most of the benefits due to very low $PM_{2.5}$ baseline exposures, but the RIA also tells us that they are very small exposure changes: the changes in exposure average only 0.7 $\mu g/m^3$ and do not exceed 1.49 $\mu g/m^3$ in any location. (RIA, p. 4-5.)

debated in terms of its implication that an implausible proportion of total deaths in the US would be due to $PM_{2.5}$ – but perhaps now it should be debated that way too.

2. Technical Gaps and Limitations Continue to Undermine the Credibility of the $PM_{2.5}$ Risk Analyses

I have not yet even mentioned the many technical aspects of the $PM_{2.5}$ epidemiological literature that also undercut the credibility of the $PM_{2.5}$ mortality co-benefits estimates for the Proposed Rule on which I am commenting. They are many, and they have been deeply explored in the comments on $PM_{2.5}$ NAAQS decisions. They include questions and debate about:

- The statistical detectability of thresholds and other forms of non-linearity in the true concentration-response relationships, to which I already alluded at the end of Section IV.1.

- Whether all fine particles are equally potent, as EPA assumes, despite their vastly different chemistries. This huge uncertainty has been identified as a problem in $PM_{2.5}$ risk analyses in multiple forums, including in National Academy of Sciences studies (such as NRC, 2002), in EPA research priority workshops, and it is recognized by EPA in this RIA.[44]

- Confounding and whether the observed associations are biased due to some other co-varying pollutant. (This bias may occur even if that co-pollutant was included in the estimation model.)

- Whether observed associations are biased due to missing explanatory variables other than co-varying pollutants (such as unidentified socioeconomic factors or locational factors such as noise from traffic). A recent and innovative paper in the *Journal of the American Statistical Association* finds that confounding appears to be playing a significant role in the statistical findings of positive $PM_{2.5}$-mortality associations (Greven *et al.*, 2011).[45]

All of these concerns are highly relevant to this particular RIA. For example, about 97% of the $PM_{2.5}$ co-benefits are due to changes in a single PM constituent, sulfate. If sulfate happens to be a non-potent constituent, then effectively all of the co-benefits in this

[44] RIA, p. 1-16 and p. 6-15.

[45] While reproducing the overall evidence of a relative risk from $PM_{2.5}$ nationwide, they find that the effect of declining $PM_{2.5}$ appears to be effectively zero when it occurs within a city rather than as part of a US-wide reduction. If there is an effect for a $PM_{2.5}$ reduction that occurs nationally, there should be a comparably-sized effect when a reduction occurs only locally. Thus, the authors express concern that they have found evidence of confounding in the $PM_{2.5}$ chronic risk associations.

particular RIA become zero, including the upper bound estimates. There is a reasonable amount of clinical evidence that suggests sulfates are not potent, yet EPA does not address this possibility in its range of co-benefits estimates. At a minimum, the lower bound of the range of co-benefits for the Proposed Rule should be zero based on this single unresolved uncertainty. Nevertheless, the total body of concerns leaves even the question of causality in the observed statistical associations between $PM_{2.5}$ and mortality still in question.

My previous comments on the $PM_{2.5}$ risk assessment being performed for the current review of the $PM_{2.5}$ NAAQS (Smith, 2009) explain much more fully why the presumption of causality remains a significant uncertainty even in the face of many papers reporting a $PM_{2.5}$-mortality risk association. For these comments, I simply summarize my earlier comments by saying that the lower bound of the $PM_{2.5}$ co-benefits should be zero. A full copy of Smith (2009) is provided in Appendix C of these comments.

3. Potential Double-Counting Across Multiple RIAs Also Undercuts the Credibility of the $PM_{2.5}$ Co-Benefits Estimates

In order for there to be *co-benefits* from $PM_{2.5}$ to attribute to the Proposed Rule, the Proposed Rule must require more reductions of primary $PM_{2.5}$ and $PM_{2.5}$ precursors (*e.g.*, SO_2 and NO_x) than would otherwise occur under other existing regulations, including the current National Ambient Air Quality Standards (NAAQS) for $PM_{2.5}$. To include any co-benefits from reductions that will occur anyway as a result of the current $PM_{2.5}$ NAAQS in this rule would be to *double*-count those benefits – first as the direct benefits that were counted to justify the $PM_{2.5}$ NAAQS in that rule's 2006 RIA (EPA, 2006), and then again as co-benefits to justify this Proposed Rule. Whether such double-counting is occurring is not clear, but it is clearly inappropriate to count as co-benefits here those attributed to $PM_{2.5}$ reductions unless those $PM_{2.5}$ reductions are strictly coincidental and incremental to what would occur anyway under the already existing $PM_{2.5}$ NAAQS and other existing standards that will be directly controlling $PM_{2.5}$ precursor emissions, such as the hourly SO_2 and hourly NO_2 NAAQS that were newly established in 2010, and the ozone NAAQS of 2008.

Some may argue that Figure 1 in Section III above demonstrates that there is very little double-counting of benefits that were originally attributed to the $PM_{2.5}$ NAAQS. The fact that Figure 1 reveals that most of the projected baseline $PM_{2.5}$ exposures are below the annual standard in 2016 should not be surprising, given that EPA's baseline scenario for 2016 includes all of the existing regulations now in place, plus the proposed Clean Air Transport Rule.[46] These include many regulations intended to create compliance with the 1997 $PM_{2.5}$ NAAQS, for which the latest attainment date is 2015.[47] Thus, it would be quite surprising if that 2016 baseline continued to show large areas of the US out of

[46] See RIA, pp. 3-14 through 3-19 for a listing of all the existing rules included in the 2016 baseline.

[47] 72 FR 20586 (Apr. 25, 2007). Two possible one-year extensions are available after 2015.

attainment with the current annual $PM_{2.5}$ NAAQS, even though EPA has stated that it did not directly simulate attainment with that NAAQS in its benefits calculations.[48] However, Figure 1 does not disprove the presence of double-counting for two reasons, the first still related to the annual $PM_{2.5}$ NAAQS, and the other related to other standards.

With regards to the annual $PM_{2.5}$ NAAQS, the estimated direct benefits reported for that standard were not limited just to changes in $PM_{2.5}$ that started above 15 $\mu g/m^3$. That analysis calculated benefits from all changes down to the LML of each epidemiological study considered. Thus, many of the benefits down to the LMLs were probably also counted back in 2006 as direct benefits from the annual $PM_{2.5}$ NAAQS (EPA, 2006), and, if so, are being double-counted now. It is impossible to determine the amount of double-counting going on in the Proposed Rule's RIA, but it could be as much as 20% of the upper-bound co-benefits estimate and 85% of the lower bound co-benefits estimate.

Double-counting concerns are not limited to those benefits that were counted when setting the annual $PM_{2.5}$ standard. EPA provides no information in the Proposed Rule, its RIA, or in any related technical support documents to indicate the extent of projected attainment with the 24-hour average $PM_{2.5}$ standard of 35 $\mu g/m^3$ in the 2016 baseline scenario. Even though the co-benefits are calculated as a function of the annual average $PM_{2.5}$ levels, some of the co-reductions being counted in this RIA may *not* be coincidental and incremental to further actions that will be needed anyway to attain the existing *daily* $PM_{2.5}$ standard, and would therefore represent double-counting of benefits that were already counted in the $PM_{2.5}$ RIA (EPA, 2006). Given that EPA has never once stated in this Proposed Rule's RIA that it has first assumed attainment of the $PM_{2.5}$ NAAQS (which includes both the annual and daily standards) before starting to count co-benefits from the MACT Rule, there is a distinct possibility that some, or perhaps a significant fraction of the 6,800 to 17,000 deaths are indeed being double-counted with respect to the daily $PM_{2.5}$ standard.

EPA should provide clear evidence that no such double-counting exists in the MACT RIA, including by showing how much of its estimated co-benefits are attributable to different levels of *daily* $PM_{2.5}$ design values in the 2016 baseline. In other words, EPA needs to provide a figure comparable to the RIA's Figure 6-15 (*i.e.*, Figure 1 above) but with the horizontal axis in units of baseline *daily* $PM_{2.5}$. Similar information should also be provided on the faction of these co-benefits attributable to days on which the new hourly NO_2 and SO_2 NAAQS are being exceeded, and the ozone NAAQS. Even though these other NAAQS have initial attainment dates after 2016 (*i.e.* 2017 and 2018, only one to two years later), there is a distinct possibility that many of these co-benefits would be merely *temporary* incremental reductions for the single year 2016, and thus would be double-counted in all but that first year of implementation of the Proposed Rule.[49]

---

[48] RIA, p. 2-11.

[49] The 2006 24-hour average $PM_{2.5}$ NAAQS may have some timing issues as well, but to a lesser degree. Its initial date for attainment is 2014, prior to the 2016 in this analysis, but extensions may be granted in some locations until 2019.

Until such additional graphs or equivalent data are provided, one cannot ascertain whether or not there is significant double-counting in the RIA's co-benefits estimates. But to the extent that double-counting exists, the co-benefits in this RIA are not credible. EPA needs to clearly demonstrate that double-counting is not occurring.

V. Summary

The lower bound of the $PM_{2.5}$ co-benefits should be zero due to the possibility that there is no causal relationship between the predicted changes in $PM_{2.5}$ and mortality. The range of reasons for this possibility to remain include the potential for systematic bias in the basis for EPA's causality inference, and the possibility that sulfates (the only $PM_{2.5}$ constituent being reduced by the Proposed Rule) have no potency among the many $PM_{2.5}$ constituents. In addition, the upper bound of co-benefits is just not plausible.

More importantly, I conclude that EPA's argument that there is a strong cost-benefit justification for the Proposed Rule is inappropriate because it is based solely on a preponderance of co-benefits from a pollutant that is already regulated, and not an air toxic. Moreover, the estimate is almost entirely derived from changes in very low concentrations that EPA has deemed adequately protect the public health. In the meantime, EPA has not been able to quantify, or even clearly identify, any meaningful amount of direct benefits from the reductions in air toxics that this rule mandates. The maximum ratio of direct benefits to costs for all three MACT groupings is 0.0006-to-1, with a net loss of about $10.9 billion per year. Each individual MACT grouping appears to impose a net benefit-cost loss on the basis of its direct benefits only, and two of those groupings appear to impose net losses even if their share of the upper bound estimates of co-benefits is included in the net benefit calculation.

References

Clean Air Scientific Advisory Committee (CASAC). 2010. Letter from J.M. Samet, Chair of CASAC, to Administrator Jackson. "CASAC Review of *Policy Assessment for the Review of the PM NAAQS – Second External Review Draft* (June 2010)" and attachments. EPA-CASAC-10-015, September 10. Available at: http://yosemite.epa.gov/sab/sabproduct.nsf/CCF9F4C0500C500F8525779D0073C593/$File/EPA-CASAC-10-015-unsigned.pdf.

Fann, N., A.D. Lamson, S.C. Anenberg, K. Wesson, D. Risley, B.J. Hubbell. 2011. "Estimating the National Public Health Burden Associated with Exposure to Ambient $PM_{2.5}$ and Ozone" accepted for publication by *Risk Analysis*, DOI:10.1111/j.1539-6924.2011.01630.x.

Greven, S., F. Dominici and S. Zeger. 2011. "An Approach to the Estimation of Chronic Air Pollution Effects Using Spatio-Temporal Information" accepted for publication by *Journal of the American Statistical Association*, DOI: 10.1198/jasa.2011.ap09392.

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

Krewski D, Jerrett M, Burnett RT, Ma R, Hughes E, Shi, Y, et al. 2009. *Extended Follow-Up and Spatial Analysis of the American Cancer Society Study Linking Particulate Air Pollution and Mortality*. HEI Research Report Number 140, Health Effects Institute, Boston, MA.

Laden, F., J. Schwartz, F.E. Speizer, D.W. Dockery. 2006. "Reduction in Fine Particulate Air Pollution and Mortality." *American Journal of Respiratory and Critical Care Medicine* 173:667-672.

National Research Council (NRC). 2002. *Estimating the Public Health Benefits of Proposed Air Pollution Regulations*. National Academies Press, Washington DC.

Pope, C.A., III, R.T. Burnett, M.J. Thun, E.E. Calle, D. Krewski, K. Ito, G.D. Thurston. 2002. "Lung Cancer, Cardiopulmonary Mortality, and Long-term Exposure to Fine Particulate Air Pollution." *Journal of the American Medical Association* 287:1132-1141.

Smith, A.E. 2009. "Comments on the External Review Draft of EPA's 'Risk Assessment to Support the Review of the PM Primary National Ambient Air Quality Standards'," prepared for Utility Air Regulatory Group, American Petroleum Institute and American Chemistry Council, submitted to PM Docket as part of UARG Comments, November 8. (A full copy is provided in Appendix C of this set of comments.)

Strum, M., J. Thurmon, M. Morris. 2011. "Non-Hg Case Study Chronic Inhalation Risk Assessment for the Utility MACT Appropriate and Necessary Analysis," Memorandum to Docket, U.S. Environmental Protection Agency, Docket No. EPA-HQ-OAR-2009-0234, March 16. Available at: http://www.epa.gov/ttn/atw/utility/pro/non-hg_risk_tsd.pdf.

U.S. Environmental Protection Agency (EPA). 2011a. *Regulatory Impact Analysis of the Proposed Toxics Rule: Final Report,* March. Available at: http://www.epa.gov/ttn/atw/utility/ria_toxics_rule.pdf.

U.S. Environmental Protection Agency (EPA). 2011b. *Policy Assessment for the Review of the Particulate Matter National Ambient Air Quality Standards*. EPA-452/R-11-003. Office of Air Quality Planning and Standards, Research Triangle Park, N.C., April. Available at: http://www.epa.gov/ttnnaaqs/standards/pm/data/20110419pmpafinal.pdf.

U.S. Environmental Protection Agency (EPA). 2010a. *Regulatory Impact Analysis for the Reconsideration of the 2008 Ozone National Ambient Air Quality Standard*, January 12. Available at: http://www.epa.gov/ttnecas1/ria.html.

U.S. Environmental Protection Agency (EPA). 2010b. *Quantitative Health Risk Assessment for Particulate Matter*. EPA-452/R-10-005. Office of Air Quality Planning and Standards, Research Triangle Park, N.C., June. Available at: http://www.epa.gov/ttn/naaqs/standards/pm/data/PM_RA_FINAL_June_2010.pdf

U.S. Environmental Protection Agency (EPA). 2009. *Integrated Science Assessment for Particulate Matter (Final Report)*. EPA-600-R-08-139F. National Center for Environmental Assessment – RTP Division. December. Available at: http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=216546.

U.S. Environmental Protection Agency (EPA). 2006. *Regulatory Impact Analysis for the 2006 National Ambient Air Quality Standards for Particle Pollution,* October 6. Available at: http://www.epa.gov/ttnecas1/ria.html.

Appendix A.

Approximate Disaggregation of Co-Benefits to Three HAPs MACT Groups for Table 1 of these Comments

There are two categories of co-benefits that need to be attributed to each of the three MACT groups: $0.6 billion of co-benefits from CO_2 and $53.4 to $139.4 billion from $PM_{2.5}$ reductions. The latter must be divided between $PM_{2.5}$ reductions due to SO_2 reductions and to primary $PM_{2.5}$ emissions reductions. Table A-1 shows the breakouts for each component that will be described in this Appendix, which sum up to the disaggregation that appears in column b of Table 1 in the main body of my comments.

Table A-1. Disaggregation of Annual Benefits Estimates by Type of Co-Benefit.

	(i) CO_2 ($ billions)	(ii) Primary $PM_{2.5}$ ($ billions)	(iii) Sulfate ($ billions)	(iv) Total ($ billions)
Mercury MACT	0.12	0.52 to 1.39	0	0.64 to 1.51
Acid Gases MACT	0.30	0.52 to 1.39	50.83 to 135.22	51.65 to 136.91
Non-Hg Metals MACT	0.18	0.52 to 1.39	0	0.70 to 1.57
Total	0.6	1.57 to 4.18	50.83 to 135.22	53 to 140

CO_2 Co-Benefits Disaggregation

The reason CO_2 is reduced as a collateral result of the Proposed Rule is mainly because of retirements of EGUs that cannot justify the overall costs of control, and perhaps some degree of shifting in the dispatch from coal-fired to natural-gas-fired units. The former is probably the dominant cause because altered dispatch will be a relatively unlikely outcome in the MACT-based scenario, because these regulations do not place any price on emissions, which is what can diminish the economic incentive to run ("dispatch") a unit. They are only mandates that they either control the HAPs to the NESHAP standard, or close. But once a unit is controlled rather than closed, its owners have a strong incentive to run it to the extent possible. The decision to shut down rather than to add the necessary controls is driven in large part by the cost of all the measures combined. I have approximated this effect by attributing the CO_2 co-benefits to each of the MACT groupings in proportion to their estimated share of EPA's estimates of the $10.9 billion program costs. EPA has provided these estimated shares in Table 25 of the Proposed Rule, and they also were used in column c of my Table 1 on page 8. Applying the same percentage breakouts as in column c of Table 1, one obtains the results in column (i) of Table A-1 above.

$PM_{2.5}$ Co-Benefits Disaggregation

The first step for disaggregating the $PM_{2.5}$ co-benefits is to determine how much of the ambient $PM_{2.5}$ reduction is due to each $PM_{2.5}$ constituent. EPA states in the RIA: "We found that, reductions in NO_x and SO_x led to significant decreases in particulate sulfate and small increases in particulate nitrate, indicating that nitrate replacement limited the nitrate reductions from NO_x decreases. Reductions in directly emitted $PM_{2.5}$ were fairly modest, providing a very small change in $PM_{2.5}$."[50] However, EPA does not offer any specific shares, other than to inform us that none of the ambient $PM_{2.5}$ reductions were attributed to NO_x emissions reductions, because EPA actually found nitrate to increase. Rather than include any offsetting $PM_{2.5}$ increases in its benefits calculation, EPA says it ignored the nitrate increases. This leaves the need to estimate what share is meant by "a very small change" in $PM_{2.5}$ due to primary $PM_{2.5}$ emissions reductions. That share of the total $PM_{2.5}$ co-benefits will then be divided among the MACT groupings according to their respective contributions to primary $PM_{2.5}$ reductions. The remainder that is due to ambient sulfate $PM_{2.5}$ is due to SO_2 reductions, which are almost entirely attributable to the acid gases MACT.

Using an approximation procedure that is similar in nature to EPA's "benefits-per-ton" (BPT) method of approximating benefits from emissions reductions, I estimate that about 97% of the co-benefits are attributable to SO_2 controls and 3% are attributable to direct $PM_{2.5}$ controls. I consider even 3% to be high because I note that in Table 27 of the Proposed Rule, EPA appears to imply that *all* of its co-benefits are due to SO_2 controls, and none to primary $PM_{2.5}$ controls.[51] I made my estimates in the following manner, using data described below:

1. Identify the average mass of each constituent in each $\mu g/m^3$ of $PM_{2.5}$ concentration in 2005.
2. Identify the total tons emitted in 2005.
3. Calculate the constituent-specific unit-mass per ton emitted (CUMPT) emitted in 2005 by dividing (1) by (2).
4. Multiply the total tons of each emitted species projected in the 2016 baseline by the CUMPT to obtain the revised constituent mix associated with the 2016 baseline ambient $PM_{2.5}$.
5. Multiply the total tons of each emitted species projected in the 2016 MACT scenario by the CUMPT to obtain the revised constituent mix associated with the 2016 MACT scenario's ambient $PM_{2.5}$.
6. Calculate how much each of the emitted species contributes to the projected change in total $PM_{2.5}$ mass from the 2016 baseline to the 2016 MACT case estimated in steps 4 and 5 above. State this as a percentage of their total incremental reduction between those two cases.

[50] RIA, p. 6-11.

[51] Proposed Rule, Table 27, p. 25077.

For step 1, I first noted that almost all of the $PM_{2.5}$ co-benefits are due to changes in annual average $PM_{2.5}$ in the Eastern U.S.[52] I therefore decided to use an annual average constituent mix representative of the East in 2005. For this, I relied on data from Figure 3-17 of the $PM_{2.5}$ Integrated Science Assessment (EPA, 2009, p. 3-57), which I have copied here as Figure A-1.

Figure A-1. Copy of Figure 3-17 of $PM_{2.5}$ Integrated Science Assessment (EPA, 2009)

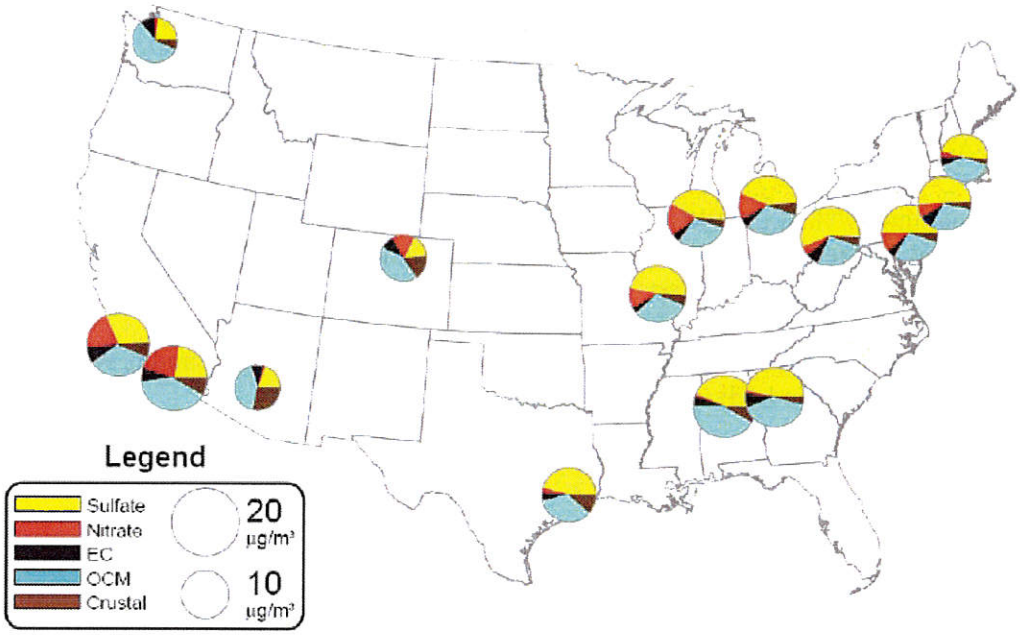


Figure 3-17. Three-yr avg $PM_{2.5}$ speciation estimates for 2005-2007 derived using the SANDWICH method. For the following 15 CSAs/CBSAs (with the number of sites per CSA/CBSA listed in parenthesis): Atlanta, GA (1); Birmingham, AL (3); Boston, MA (4); Chicago, IL (7); Denver, CO (2); Detroit, MI (4); Houston, TX (1); Los Angeles, CA (1); New York City, NY (7); Philadelphia, PA (6); Phoenix, AZ (2); Pittsburgh, PA (4); Riverside, CA (1); Seattle, WA (4); and St. Louis, MO (3). SO_4^{2-} and NO_3^- estimates include NH_4^+ and particle bound water and the circle area is scaled in proportion to FRM $PM_{2.5}$ mass as indicated in the legend.

Based on the information in Figure A-1 for Eastern sites, I assumed a constituent mix of 50% sulfate, 30% organics (OCM), 10% nitrate, and 10% EC+crustal. I combined EC and crustal ambient $PM_{2.5}$ because they are both derived from the emissions reported as primary $PM_{2.5}$. EGU primary particle emissions also are a combination of the two, although they are more significantly made up of crustal elements than EC at most modern units.

[52] See Table 1-3 of the RIA, pp. 1-5 to 1-6.

For tons emitted required in steps 2, 4, and 5, I relied on data in Table 4-2 of the RIA. Table A-2 below shows these inputs and the results of the calculations that were made using them. As one can see from the column (i) of Table A-2, my calculations suggest that reduction in ambient sulfate accounts for about 97% of the reduction in $PM_{2.5}$ mass when going from the 2016 baseline to the 2016 MACT scenario, and that EC+crustal ambient $PM_{2.5}$ accounts for about 2.5%. Since this approximation analysis does not account for the small rebound effect in nitrate, I rounded the primary $PM_{2.5}$ share up to 3% for my calculations.

Thus I attribute 3% of the $52.4 to $139.4 billion of $PM_{2.5}$ co-benefits to primary $PM_{2.5}$ controls under the Proposed Rule, for a total of $1.6 billion to $4.2 billion. I then assign this reduction in equal shares to each of the MACT groupings, since each of them to some extent results in added particulate controls. These results appear in column (ii) of Table A-1 above.

The remaining 97% of the $PM_{2.5}$ co-benefits, $50.83 billion to $135.22 billion, is from incremental SO_2 reductions, which I attribute to the acid gases MACT. This appears in column (iii) of Table A-1 above.

Column (iv) of Table A-1 shows the totals for each MACT grouping, and is what appears in Table 1 of the main body of my comments.

Table A-2. Approximate Estimation of Portion of Total $PM_{2.5}$ Co-Benefits Due to Each Emitted Species

$PM_{2.5}$ Constituent	(a) Avg quantity constituent per 1 $\mu g/m^3$ mass	(b) Tons emitted in 2005	(c) CUMPT (per million tons)	(d) Tons emitted in 2016 baseline (millions)	(e) Avg quantity constituent per $\mu g/m^3$ 2005 mass remaining in 2016 baseline	(f) Tons emitted in 2016 MACT scenario	(g) Avg quantity constituent per 1 $\mu g/m^3$ 2005 mass remaining in 2016 baseline	(h) reduction in mass 2016 due to MACT rule compared to baseline	(i) %PM reduction due to each species from 2016 baseline to 2016 MACT
Sulfate	0.5	15.05	0.033223	7.25	0.240864	4.89	0.162458	0.078405	96.6%
OCM	0.3	17.6	0.017045	14.39	0.245284	14.39	0.245284	0	0.0%
Nitrate	0.1	22.2	0.004505	15.02	0.067658	14.87	0.066982	0.000676	0.8%
EC+crustal	0.1	4.4	0.022727	4.02	0.091364	3.93	0.089318	0.002045	2.5%
TOTAL MASS	1.0				0.645169		0.564043	0.081126	

(a): Input based on data in Figure 3-17 of EPA (2009) – step 1.
(b): Input based on data in RIA Table 4-2 – step 2.
(c): Calculated: column a/column b – step 3.
(d): Input based on data in RIA Table 4-2 – step 4.1.
(e): Calculated: column c * column d – step 4.2.
(f): Input based on data in RIA Table 4-2 – step 5.1
(g): Calculated: column c * column f – step 5.2.
(h): Calculated: column e – column g – step 6.1
(i): Calculated: column h/total for column h – step 6.2.

Appendix B.

Brief Overview of How EPA Calculates its $PM_{2.5}$ Mortality and Morbidity Risk Estimates Based on Epidemiological Studies

EPA's own documents provide a thorough explanation of its risk calculation methods. The purpose of this appendix is to provide a brief overview of EPA's steps that is aimed less at comprehensive documentation and more at providing insight about the underlying assumptions and their impact. It is provided merely as background, and to help explain some of my own inferences in the main body of my comments.

The health effects estimates, whether they be mortality or morbidity endpoints, are created through two totally separate steps: (1) epidemiology and (2) risk analysis. The risk analysis uses a number of parameters, only one of which is taken from the results of an epidemiological study. That input parameter from the epidemiological literature happens to be one of the most critical parameters in the risk analysis, but it is not the only critical risk analysis assumption. However, the issues and concerns in doing an epidemiological study and the risk analysis are quite different and I will summarize each separately.

The Epidemiological Literature for $PM_{2.5}$

Epidemiology, which is the study of health-related events and patterns in populations, involves many different types of research, and I will not attempt to list them all. The most prominent type of epidemiological study used to form inputs for EPA's $PM_{2.5}$ risk analyses are statistical analyses of the correlation between monitored ambient $PM_{2.5}$ and a health outcome such as dying, or entering the hospital with a heart attack. These statistical analyses are done with many more control variables than simply $PM_{2.5}$, (that is, they are multivariate in nature). Their primary goal is to report how much more of a particular type of health outcome tended to occur when $PM_{2.5}$ monitors had higher readings than when they had lower readings and which cannot be attributed to changes in any other health-affecting variables. Thus, the epidemiological findings are, at heart, a statistical correlation.

The greatest challenge in these epidemiological studies revolves around finding the most appropriate set of other variables that are also likely to be explaining the risk of the health outcomes being studied, and also finding data sets that properly reflect those other risk factors for the population whose health outcomes are being studied. Without proper controls for all the relevant risk factors, the $PM_{2.5}$ variable might take on some of the explanatory role for the "missing variables." This can occur if the missing variable and the $PM_{2.5}$ data happen to have some correlation with each other.

In almost all of the $PM_{2.5}$ epidemiological studies, the variable explaining a population's exposures to $PM_{2.5}$ is based on population-centric $PM_{2.5}$ monitors, most often from the EPA AIRS network. This means that thousands to hundreds of thousands of people are

all assumed to be breathing $PM_{2.5}$ at the exact same level as reported by the nearest monitor to their home or at a multi-monitor average level computed for their entire urban area. The $PM_{2.5}$ epidemiological studies that are used in EPA's risk analyses never know exactly what $PM_{2.5}$ level individuals in the study were breathing, so the studies are working with very approximate data on exposures.

Similarly, these studies are relying on very approximate data about the relevance of the population's health outcomes to ambient $PM_{2.5}$. By this I mean that when a person is observed to have died in a study that follows the health outcomes of specific individuals, there is no information to identify that death as $PM_{2.5}$-related, or even as pollution-related. There is only information on the day and age of death and the cause(s) of death stated on the death certificate. Causes of death are reported using a standard code that identifies events such as stroke, heart attack, asthma, accidental, etc. There is no code for "pollution."

There are two basic types of these $PM_{2.5}$ studies: acute and chronic exposure studies. EPA's risk analyses for mortality nowadays rely entirely on the chronic types of studies, but it is important to understand the acute studies approach for two reasons. First, many of the morbidity endpoints in the RIA are based on acute studies. Second, EPA used to report mortality risks from both acute and chronic studies, although it has in recent years (including in the Proposed Rule's RIA) projected mortality risks using chronic studies only. In the case of $PM_{2.5}$, the specific chronic mortality risk estimates that EPA relies on have always been much larger than any of the available acute mortality risk estimates.

Acute studies track short-term (*e.g.*, daily) changes in air pollution and correlate these changes with short-term variations in health effects, while controlling for other time-varying health-affecting factors too. Because most short-term variations in air pollution are best characterized locally, acute studies are, at their foundation, based on specific locations (*e.g.*, individual cities), and are "time-series" studies in nature.[53] Many acute studies, and particularly those for mortality and hospitalizations, do not follow any specific group of identifiable individuals, but instead consider the entire population of a city or Standard Metropolitan Statistical Area (SMSA). They do so by correlating the city's or SMSA's officially reported numbers of deaths (or other health-related event, such as hospitalizations) on each day against that city's monitored $PM_{2.5}$ on that day.[54]

---

[53] The most sophisticated of the acute $PM_{2.5}$ studies actually consider multiple cities, and are referred to as "multi-city studies." However the fundamental analysis occurs at the level of each city individually, as I describe in the rest of this paragraph. More complicated statistical techniques are layered on to develop greater statistical power by considering commonalities in the observed associations across different populations. I will not go into the details of multi-city methods, as my purpose here is to explain just the basic concepts of the epidemiological methods being applied to $PM_{2.5}$.

[54] Acute studies of health events that are not tracked or reported in official public health data sets, such as wheezing or days of restricted activity, are done by identifying a cohort of individuals in a given location and having them self-report their own experiences of these events in order to create the necessary data base for statistical analysis of how these events may be correlated with short-term exposures to $PM_{2.5}$. These types of acute epidemiological studies are often called "panel studies."

As in any epidemiological study, the major challenge of the analysis is to properly control for all the other important variables that explain why deaths (or other health effects) are higher on some days of the year than on others. These variables can include, among others, season, temperature, humidity, day of week, and presence of known epidemics. Once the researchers feel they have captured as many of the non-$PM_{2.5}$ explanatory factors as appropriate in the equation that they are statistically estimating (this equation is often called "the model"), they include the city's measured $PM_{2.5}$ and statistically determine how much of the remaining ups and downs in daily reported events coincides with the concurrent ups and downs in monitored ambient $PM_{2.5}$ for the entire area. Often the researcher lets the model account for some degree of lag in an observed correlation, which is interpreted as saying that the health effect does not manifest itself until one or more days after the $PM_{2.5}$ exposure occurs.

These studies are called "acute" effects studies because they only capture increases in the studied health effect that might be resulting from short-term and relatively recent increases in ambient $PM_{2.5}$. When an acute study reports an association with mortality, it is commonly assumed that this mortality association, if causal at all, is affecting individuals who are already in a poor state of health.

Chronic studies contrast to acute studies in their fundamental emphasis on estimating differences in average health risks across many locations whose long-term pollution levels (*e.g.*, annual average concentrations) differ.[55] Because the differences in locational health effects rates are tied to long-term average levels of exposure, these are also called "chronic" risk studies. That is, they are essentially cross-sectional rather than time-series studies. The best of the chronic studies (and the only kind that EPA relies on for inputs to its risk analyses) compare the health effects incidence differences for a well-identified set of specific individuals, rather than just population-wide average incidence rates. This allows for the researchers to account for some of the specific differences in the individuals that might also affect their health outcomes.

For *chronic morbidity studies*, researchers must contact the individuals they have recruited into the cohort at least once again (at the end of the study period) because morbidity outcomes cannot be tied to specific individuals through public data sets, and therefore the individuals must report their personal health outcomes to the researchers directly. Also, because a person can experience non-fatal health events on multiple occasions (whereas death only occurs once), enough data to correlate differences in incidence rates with differences in locational air pollution levels may be possible to obtain in only a few years from cohort initiation. Chronic morbidity studies therefore can

[55] For example, of the chronic mortality studies that are mentioned in the RIA, Pope *et al.* (2002) and Krewski *et al.* (2009) are both based on a cohort called the American Cancer Society or "ACS" cohort, the individuals studied are located in dozens of cities. Laden *et al.* (2006), which provides the upper bound mortality risk estimate in the RIA, is based on individuals located in only six cities in the US, hence this is called the "Six Cities" cohort. The Medicare cohort that has been studied by Greven *et al.* (2011) includes individuals in dozens of cities.

have time spans that make it viable to stay in contact with a sufficient fraction of the individuals in the cohort.

Chronic mortality studies, in contrast, can require longer periods of time before a sufficient number of deaths have occurred to correlate differences in mortality risk levels with locational differences in air pollution levels. Thus, the survival outcomes of individuals in cohorts used for chronic mortality risk studies are often observed many years after their recruitment. Maintaining follow-up contact with a sufficient fraction of the cohort over such a long period after recruitment can be a challenge, but an offsetting advantage for chronic mortality studies is that individuals' deaths are reported in public records such as death certificates. Thus, researchers can develop a data base of cohort survival outcomes without ever contacting the individuals again after recruitment. Thus, in many of the chronic mortality studies, individual-level data (*e.g.*, sex, race, health conditions at time of recruitment, smoking behavior, education, residential location) may not be updated after the time of their initial recruitment. Lack of tracking of the individuals in some mortality cohorts can create uncertainty in the quality of the individual-level data used as control variables because some of the important factors determining individuals' survival outcome (such as smoking behavior and where exactly they are living) can change in the many years that may pass between their recruitment and their death.

In the RIA, the $PM_{2.5}$ co-benefits are almost entirely due to chronic mortality risk. Thus, the rest of this discussion focuses solely on how mortality risk studies are performed and used. Over many years of tracking a cohort through death certificate records, researchers studying a mortality risk cohort are able to build up a "survival curve" for the set of people in the cohort residing in each location (the location may be defined as a city, SMSA, county, or even zip code). A survival curve is an actuarial concept that reports for a defined population the likelihood that a random person of any age within that population will survive into the next year of his/her life.[56] The cohort may be followed for decades.[57] The statistical analysis that is applied to the chronic mortality study's data set strives to explain differences in the survival curves observed for each location as a function of information about each individual's age, other personal characteristics, and information about the location's average socio-economic characteristics (used when the comparable individual-level data are not available). As noted above, in many cases, the individual-level data may have only been obtained at the time of recruitment, including where the person is assumed to live.

[56] Related actuarial information that can be derived from a survival curve is the probability of dying this year given one's age (*i.e.*, the age-specific population mortality rate), and life expectancy given one's current age. Overall population mortality rates and life-expectancy can also be derived from a survival curve.

[57] For example, one of the most widely cited prospective cohorts, the ACS cohort used by Pope *et al.* (2002) was first recruited in 1982 and their survival outcomes continue to be followed today. This explains why there are periodic "update" studies for the chronic risks: as more years of data are collected, there are more deaths in the cohort, which helps the researchers obtain a more and more statistically-sound representation of that population's actual survival curve – particularly at the higher age levels.

Once researchers feel they have explained the differences in survival curves of the cohorts in each location as well as possible with all the available data for non-pollutant variables, they statistically estimate how much of the remaining unexplained differences in the survival curves are correlated with differences in the locations' average ambient concentrations of $PM_{2.5}$. As with the acute studies, a single $PM_{2.5}$ exposure is assumed for every person within a location. For example, when the location is defined as a county, every person in the cohort who is believed (based on their reported domicile at time of recruitment into the cohort) to be living within that county is assumed to be exposed to the same ambient $PM_{2.5}$ concentrations recorded at that county's monitoring station (or composite of that country's monitoring stations, if more than one exist). The statistical method that is most commonly used (called the "Cox Proportional Hazards" method) assumes *by construction* that $PM_{2.5}$ causes the survival curve to shift by the same proportion at all age levels. The resulting statistical estimate of difference in mortality risk per unit of difference in $PM_{2.5}$ among all the locations in the study is therefore a single mortality risk that is applied to all ages.[58]

Since chronic mortality studies follow individuals over many years, it becomes an interesting question as to which average $PM_{2.5}$ concentration to use "in the model" for the statistical correlation, given that ambient $PM_{2.5}$ concentrations have been changing over that same period. They have been generally declining, but not by the same amount in each city. The choice of year(s) to correlate survival outcomes for the $PM_{2.5}$ data affects the statistically-estimated parameter that quantifies the amount of risk associated with each unit of $PM_{2.5}$ increase, even though the same set of deaths are being explained.[59] There is no right answer, given the very broad patterns of mortality risk that are being studied and lack of any clinical understanding of how $PM_{2.5}$ exposures actually affect one's health, but the choice made by the researchers, or the choice that EPA makes for its risk analysis when relying on a study that offers parameter estimates from several different temporal averages of $PM_{2.5}$, can greatly affect the premature mortality risk estimates that appear in an RIA.

---

[58] When risk analyses are performed using this single mortality risk for all age groups, the numbers of deaths that are calculated do, nevertheless, concentrate among the elderly. This is simply because there is a much greater baseline mortality rate among the elderly, and so when the same proportional risk change is applied to their much higher baseline mortality rates, more premature deaths will be projected in those older age groups. This fact does not, however, imply that the original epidemiological study identified (or even could have identified) that the *relative* risk was greater at older ages.

[59] For example, if the same set of cohort deaths (summarized as locational survival curve differences) is statistically correlated with $PM_{2.5}$ from an earlier period when it was higher and then with $PM_{2.5}$ from a later period when it was lower, the estimated change in mortality risk per unit of $PM_{2.5}$ change will be smaller for the earlier period and larger for the later period. Some might interpret this as implying that estimates of risks are higher with the more recent $PM_{2.5}$ data, when in fact it is strictly a result of explaining a single fixed amount of change in risk using data on $PM_{2.5}$ that contain different absolute amounts of locational differences in $PM_{2.5}$. See pp. 8-9 in Appendix C of my comments for an example that indicates this is the case.

"Relative Risk" is the key summary statistic from the epidemiological studies that EPA reports in the RIA. Relative risk states how much the risk of dying at any age is increased per unit increase in $PM_{2.5}$. Most often today, the relative risk is stated per 10 $\mu g/m^3$ of $PM_{2.5}$. For example, one of the mortality relative risk results in Laden *et al.* (2006) – the one that EPA chooses to use for its upper bound mortality risk estimate in the Proposed Rule's RIA – is 1.16 per 10 $\mu g/m^3$.[60] This should be interpreted as saying that people in a city with, say, 20 $\mu g/m^3$ annual average $PM_{2.5}$, are estimated to be experiencing a 16% higher mortality rate than people in a city with 10 $\mu g/m^3$ annual average $PM_{2.5}$.[61]

The relative risk estimate is an average increase that has been statistically inferred from a range of actually observed $PM_{2.5}$ data that has a lowest measured level (LML) and a highest measured level (HML). Any risk analysis that calculates risk for changes in $PM_{2.5}$ that occur below the LML or above the HML is extrapolating beyond the data on which the relative risk estimate was based. Such extrapolation is subject to great uncertainty. Indeed, if the available observations in the epidemiological data set are very sparse just above the LML or just below the HML, then there is also substantial uncertainty with risk estimates made for exposures in those infra-marginal concentration ranges: greater uncertainty than is implied by the reported statistical confidence interval.[62] The LML that is reported for the Laden *et al.* (2006) study is 10 $\mu g/m^3$ annual average $PM_{2.5}$ (see Figure 1 in the main body of my comments). However, that is the LML for the most recently *estimated* $PM_{2.5}$ levels in that dataset. In the earlier years' *observed* ambient averages for that same study, the LML was 11.4 $\mu g/m^3$.[63] Similarly, while EPA is stating that the LML for the Pope *et al.* (2002) study is 7.5 $\mu g/m^3$, the LML for that same cohort in its earlier years (1979-1983) was greater than 9.0 $\mu g/m^3$.[64] The problem with assuming that increased relative risks have been observed for $PM_{2.5}$ concentrations that extend to as low as 10 $\mu g/m^3$ in the case of Laden, or 7.5 $\mu g/m^3$ in the

---

[60] Laden *et al.* (2006), p. 670.

[61] One should be aware that the relative risk is sometimes reported per 1 $\mu g/m^3$ or even, in some of the earlier studies, by unusual numbers like "per 24.5 $\mu g/m^3$" (*e.g.*, in Krewski *et al.*, 2000). It is therefore important to make sure that one knows what increment of $PM_{2.5}$ is being assumed when a relative risk is reported, and never to compare relative risks that are stated for different $PM_{2.5}$ increments. When stated per 1 $\mu g/m^3$, this is a close approximation of the actual parameter estimated in the study, which is called either the "risk coefficient" or the "beta coefficient." The latter parameter, not the relative risk value, is what is actually used in the risk analysis calculations, as will be explained in the next section.

[62] CASAC discussed these points in its letter to EPA reviewing the Second Draft of the $PM_{2.5}$ Policy Assessment Document (CASAC, 2010, pp. 2-5).

[63] I emphasize *estimated* and *observed* because Laden *et al.* (2006) performed their update of the Six-Cities cohort risk analysis, they did not use actual monitored $PM_{2.5}$ for the more recent time period's concentration data, but rather estimated those data. The earlier $PM_{2.5}$ data set for that cohort was actually monitored (Laden *et al.*, 2006, p. 668). The point I am making in this paragraph, however, would be the same even if all of the concentration data had been observed.

[64] Krewski *et al.* (2000), Table 1, p. 35. (The LML was higher than 9.0 $\mu g/m^3$, because the earlier dataset reported median concentrations, whereas the later analyses are relying on mean concentrations, which are higher than medians.)

case of Pope is that no one knows whether the increased risk being associated with $PM_{2.5}$, *if causal*, is attributable to the earlier and higher $PM_{2.5}$ exposures that these same people experienced, or to the more recently estimated levels. Thus, the appropriate ambient concentration for where uncertainty in the statistically-estimated relative risk parameter begins to be a dominant concern is nebulous, even when an LML is reported. The LML associated with the most recent ambient concentrations is the *lowest* of the concentrations that can be considered appropriate to use, and there are concerns with using even that LML.

The important point, however, is that while the statistically-estimated relative risk of 1.16 might imply that mortality risk is 16% higher for a city that is at 20 $\mu g/m^3$ compared to in a city that is at 10 $\mu g/m^3$ (*i.e.*, for cities that fall within the range of concentrations actually studied), it is not appropriate to also say that a city with $PM_{2.5}$ at 10 $\mu g/m^3$ is experiencing risk that is 16% higher than if it could reduce its $PM_{2.5}$ to 0 $\mu g/m^3$ because the studies have not been able to statistically observe risks in that lower exposure range.

A comparable numerical example that is closer to the actual calculations that are going on in the RIA is as follows. It is appropriate to summarize the statistical findings of Laden *et al.* as suggesting that the risk in a city with ambient $PM_{2.5}$ of 16 $\mu g/m^3$ might be experiencing overall mortality risk that is about 1.6% higher than a city with ambient $PM_{2.5}$ of 15 $\mu g/m^3$ (*i.e.,* these are concentrations well within the observed levels in the study, and a 16% risk increase per 10 $\mu g/m^3$ also implies – *approximately* – a 1.6% risk increase per 1 $\mu g/m^3$.) However, that same study cannot be said to have shown that a city with ambient $PM_{2.5}$ of 10 $\mu g/m^3$ might be experiencing overall mortality risk that is 1.6% higher than a city with ambient $PM_{2.5}$ of 9 $\mu g/m^3$ because differences in concentrations in this lower range have not actually been observed by that study. In other words, those changes in concentration are below that study's LML. The RIA risk analysis is making the latter type of assumption, however. That is, 70% of the RIA's estimates of mortality risk are for $PM_{2.5}$ below about 10 $\mu g/m^3$ (see Figure 1 of my comments), and the average change in $PM_{2.5}$ is 0.7 $\mu g/m^3$.[65] Extrapolation of risks below the LML is thus an inappropriate use of statistically-derived relative risk values that produces risk estimates not based on any public health data or other evidence.

The Risk Analysis Method EPA Uses for $PM_{2.5}$

The risk analysis is strictly a calculation. Its only basis in empirical data on $PM_{2.5}$ risks is that it relies of the relative risk estimates from one or more of the epidemiological studies.[66] Thus, while the epidemiological literature is focused on estimating what the relative risk might be based on empirical data, the risk analysis is focused on calculating

[65] RIA, p. 4-5.

[66] EPA also frequently mentions its "expert elicitation" as an additional basis for its relative risk assumptions, but these are merely experts' opinions about the true relative risk based on their knowledge of the many different relative risk values reported in the $PM_{2.5}$ epidemiological literature.

what those statistically-derived relative risk estimates, *if assumed to reflect a causal relationship*, imply about changes in health outcomes if ambient $PM_{2.5}$ concentrations were to be changed.

In order to make risk calculations, one has to first decide whether the observed statistical associations in the epidemiological literature represent a *causal* association. EPA has concluded that these observed statistical associations are causal (PM, 2006). Whether one accepts EPA's conclusion or not, the point here is that the causal nature of those statistical associations is *presumed* in EPA's risk analyses.[67] EPA does not incorporate any uncertainty or likelihood about causality in any of its risk analysis calculations.

Given the presumption that the epidemiological associations are causal in nature, one still has to make judgments about the specific shape of the mathematical relationship between $PM_{2.5}$ and increased health risk (whether for mortality or some morbidity endpoint). In the field of toxicology, it is widely agreed that one should expect some sort of S-shaped curve for any toxic compound affecting the health of a population.[68] No one has suggested that ambient $PM_{2.5}$ levels are anywhere near the upper reaches of such an S-curve, so the main issue in defining a specific mathematical relationship to calculate how ambient $PM_{2.5}$ exposures alter population health risk is whether it has any of the non-linearity of the sort associated with the lower end of an S-curve. This is not just a question of whether there is some lower bound of $PM_{2.5}$ at which there is no relative risk; it is also an important question whether the average relative risk that is reported out of the epidemiological study that studied an entire range of observed $PM_{2.5}$ levels might be an average of a higher relative risk at the upper ends of the observed range and a lower relative risk at the lower ends of the observed range. Based on some studies of this question in the epidemiological literature, EPA assumes in its risk analysis that it will use the identical shape for its "concentration-response" as the function that was used for the statistical derivation of the relative risk. This is often described as a linear relationship, but although it is close to linear, it is actually based on a non-linear formula, *i.e.*, an exponential. The specific formula is as follows, which for purposes of explanation, I have stated in two steps:

$$\text{Relative Mortality Risk} = e^{\{(\text{change in PM2.5 in } \mu g/m3)_x(\text{risk coefficient})\}} \quad [1]$$

$$\%\ \text{Change in mortality risk} = (\text{Relative Mortality Risk} - 1) \times 100 \quad [2]$$

[67] For a discussion and critique of the basis for EPA's determination that the epidemiological evidence is causal, see Smith (2009), reproduced in Appendix C of my comments.

[68] That is, at very low exposures approaching zero, few if any people will be affected, but an increasingly large fraction of people will be affected by exposures at increasingly higher levels. After some quite extreme level, all of the population that will be affected will have been affected, and then further, higher exposures will have little incremental effect on population risk, so the exposure-response curve flattens again.

The numbers of deaths attributable to the change in $PM_{2.5}$ concentration that is input to equation [1] is calculated by simply multiplying the number of current deaths in a location by the estimated percentage change in risk that is output from equation [2].[69] EPA performs this calculation at the county level in its BenMAP model, estimating numbers of deaths reduced in each county, using county-specific estimates of changes in $PM_{2.5}$ concentrations and current numbers of deaths in that county, then sums them all up for the national estimate. As I will show below, one can make a fairly reasonable back-of-the-envelope estimate of the national total number of avoided deaths just by using an estimate of the national average of how much $PM_{2.5}$ concentrations are projected to be changed by a policy.[70]

The "risk coefficient" in the above equation is the exact numerical parameter estimated by the epidemiological "model" being relied upon for the risk analysis, rather than the "relative risk" that is usually what is reported. It is directly related to the relative risk, however, and one can easily back out its approximate value from a reported relative risk in the following way: by finding the value of the risk coefficient that when plugged into the above equation, along with the change in $PM_{2.5}$ that it is based on, produces the relative risk. For example, consider the Laden-based relative risk of 1.16 per 10 $\mu g/m^3$. The risk coefficient that goes into the equation [1] is the value of "β" that causes $\exp(10\ \beta)$ to equal 1.16. A little bit of algebra tells us that $\beta = \ln(1.16)/10$, or $\beta = 0.0148$ for the Laden-based calculations.[71]

Although the formula involves the use of an exponential form, for the ranges of $PM_{2.5}$ concentrations and risk coefficients that will be used in any risk analysis, this risk formula is nearly linear. For example, Figure B-2(A) shows, for a risk coefficient of .0058 (consistent with a relative risk per 10 $\mu g/m^3$ of 1.06), the percentage increase in mortality rates assumed for $PM_{2.5}$ concentrations ranging from 0 to 25 $\mu g/m^3$ using the

---

[69] EPA just uses current numbers of deaths per year in the location in question, and assumes the relative risk changes occur from that level. This particular assumption is innocuous in terms of numerical estimates of changes in premature mortality that result from the risk analysis.

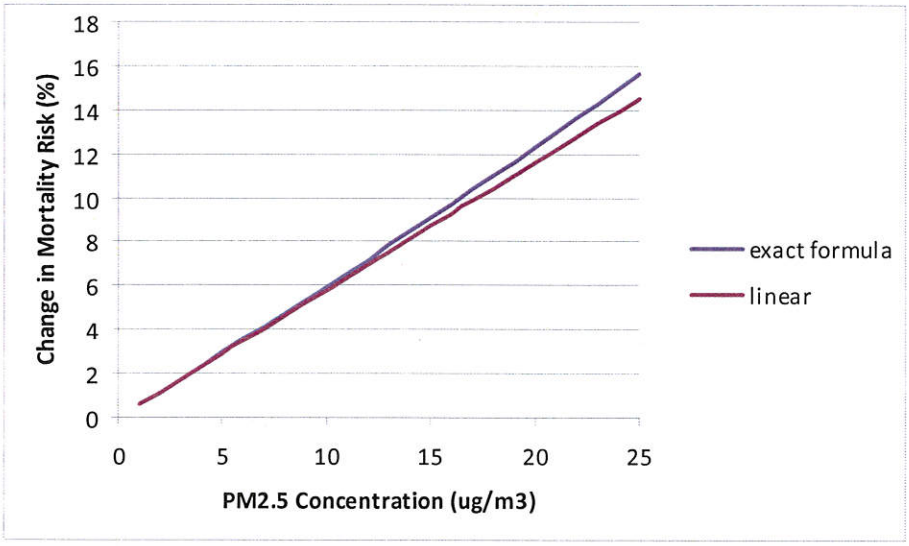
[70] A further step in the calculations appears in Appendix C of the RIA, and which I therefore also use in my comments on those estimates in Section IV.1. The calculations described so far in this appendix generate estimates of the *change* in mortality for a "before" and an "after" estimate of $PM_{2.5}$, either as a percentage change or a change in numbers of deaths. However, Appendix C of the RIA actually states that change as the "percent of total mortality that is due to $PM_{2.5}$." This requires that the percentage change in risk, R, be restated as a percent to the total deaths at that point in time, which is (1+R). Thus, whenever the mortality impacts are stated as a percentage of deaths *due to* $PM_{2.5}$, it has been calculated by first estimating R, the fractional increase from background, then reporting the value R/(1+R). For example, using a risk coefficient of .015 and a change in $PM_{2.5}$ of 20, the percent increase from that change of 20 above background is exp(20*.015)-1, which is 35%. However, that 35% increase implies that 26% of all deaths are *due to* that increase, because (0.35/1.35)=.26, or 26%.

[71] I have used the symbol β in this example because this is the symbol that EPA uses in its documentation, and when the direct estimate of the risk coefficient is reported in an epidemiological paper, it is often referred to as "β" there as well.

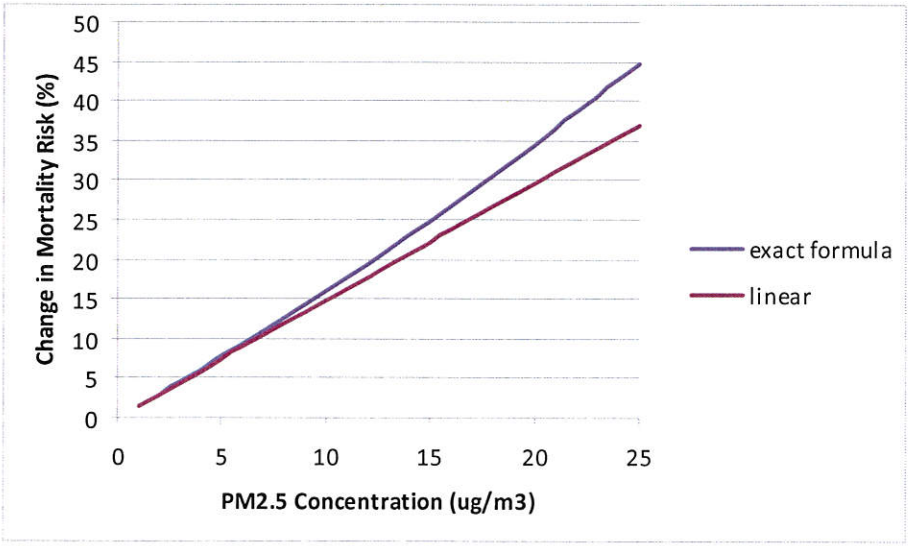
precise formulas of equations [1] and [2], and also using a linear formula that simply applies the risk coefficient times the $PM_{2.5}$ level.

Figure B-1. Comparison of Increased Risk Using Exact Exponential Formula to Linear Approximation.

(A) Using risk coefficient of 0.0058 (*i.e.*, a relative risk of 1.06 per 10 μg/m^3)



(B) Using a risk coefficient of 0.0148 (*i.e.*, a relative risk of 1.16 per 10 μg/m^3)



The exact formula's risk estimate only diverges from the linear risk estimate for $PM_{2.5}$ concentrations near the upper ranges. Figure B-2(B) provides the same comparison for a risk coefficient of .0148 (consistent with the upper bound relative risk of 1.16 per 10 $\mu g/m^3$). This is among the highest of the relative risk estimates that have been reported in the epidemiological literature, and the non-linearity of the relationship does become more pronounced. However, what is most pronounced is the much higher percentage increase in total mortality risk that is associated with this relative risk of 1.16, which can be seen by comparing the range of values on the vertical axes of the top and bottom graphs in Figure B-1.

The mathematical formulas can appear daunting, but what one should know is that a very close approximation of the risk coefficient is the increased risk per 1 $\mu g/m^3$ implied by the relative risk. For example, the Laden-based relative risk of 1.16 implies a 16% risk increase for 10 $\mu g/m^3$, which is *approximately* a 1.6% increase for 1 $\mu g/m^3$, which is 0.016. Note how close this is to the more precise estimate of β =0.0148. The ability to approximate this way is even better for the smaller relative risks such as the Pope-based relative risk of 1.06. Stated as increased risk for 1 $\mu g/m^3$, it is 0.6%, or .006, whereas the precise β for a relative risk of exactly 1.06 is 0.0058.

The practical implication of all this is that one can do very quick mental estimates of the number of premature deaths that will be calculated for a reported change in $PM_{2.5}$ without resorting to the exact risk analysis formulas that EPA uses. For example, EPA is using a relative risk of 1.06 per 10 $\mu g/m^3$ for its lower bound estimate and EPA also reports that the average reduction in $PM_{2.5}$ is 0.7 $\mu g/m^3$. With just this information, one can quickly approximate the change in premature mortality as 0.7 times the approximate 0.6% increased risk estimated per 1 $\mu g/m^3$. That is, the mortality risk would decrease by approximately 0.42% for a decrease of 0.7 $\mu g/m^3$ in annual average $PM_{2.5}$ concentration. Given that there are about 2.4 million deaths per year in the US, 0.42% implies about 10,000 deaths per year, which is not exactly the 6,800 deaths that EPA reports for its lower bound, but it was much quicker to compute than obtaining the BenMAP model and all the massive data that goes into it, then performing the benefits-per-ton estimate, to come up with the "precise" estimate of 6,800 deaths. This kind of approximation can be useful for doing "reality checks" on the EPA estimates.

Another simple back-of-the-envelope calculation is to make approximations of the numbers of deaths that would be estimated if different relative risks were to be used, which is something I have done quite extensively for my comments (see Section IV.1 of the main text of these comments). Quite reasonable approximations can be developed by simply scaling the numbers of deaths reported from calculations using one relative risk upwards or downwards by the percent change between the relative risk used originally and the alternative one. For example, EPA's estimate of 6,800 deaths is based on a relative risk of about 1.06. The upper bound is based on Laden, which has a relative risk of about 1.16, which implies about 2.67 times more increased risk per unit of change in $PM_{2.5}$. Without actually going back to BenMAP and calculating new benefits-per-ton, one can quickly estimate that the upper bound estimate will be 2.67 times 6,800, or

18,000. This is close to the 17,000 that EPA has reported, but an even more precise back-of-envelope estimate would apply the scaling up using the ratio of the risk coefficients, not the relative risks. Because the risk coefficients can be slightly different than the relative risk, the exact scaling factor can differ. For example, for a relative risk of 6% and 16%, respectively, the ratio of their risk coefficients would be (.0148/.0058), or 2.55 rather than 2.67. When 6,800 is multiplied by 2.55 and rounded to the nearest 1000, the result is 17,000, which is exactly the value EPA reports for its upper bound. I have used this more precise method with the specific risk coefficients in the calculations I have reported in the main body of my comments.

The more precise calculations that EPA performs using BenMAP are, of course, necessary to give the risk analysis reproducibility and documentability. Unfortunately there are some less desirable consequences that arise from the fact that EPA does its risk analysis calculations with so much greater detail and precision. The first is that it creates an aura of scientific sophistication and precision about the resulting risk estimates that is unfounded, because they are actually the result of some very gross assumptions about the *quantitative* meaningfulness of the epidemiological results as a specific and complete representation of an actual concentration-response function. The second, less desirable consequence is that many readers of the RIAs that report the output of the risk analyses may feel utterly incapable of understanding how EPA derived its findings, and therefore unable to review or comment on them.

My purpose in writing this background discussion has been to try to help people realize that they can roughly reproduce EPA's risk estimates, and perform their own approximate sensitivity analyses on them with little more than a hand calculator, and some of the basic inputs EPA has used, such as (a) the relative risks used, (b) the range of $PM_{2.5}$ concentrations in the baseline and (c) the changes in $PM_{2.5}$ concentrations. This capability is important for reviewers who wish to make useful comments on the RIAs and risk analyses that EPA produces. I hope that with the benefit of the information in this appendix, more people will feel competent to understand the basis of EPA's mortality and morbidity estimates, and to draw their own conclusions about them.

Appendix C.

Full Copy of Smith (2009) Comments on the Weakness of the Case that $PM_{2.5}$ Has a Causal Relationship with Mortality Risk.

CRA Charles River Associates

Comments on the External Review Draft of EPA's "Risk Assessment to Support the Review of the PM Primary National Ambient Air Quality Standards"

Anne E. Smith, Ph. D.
Charles River Associates
1201 F Street NW, Suite 700
Washington, DC 20004

Prepared at the request of the Utility Air Regulatory Group, American Petroleum Institute and American Chemistry Council

November 8, 2009

In September 2009, the U.S. Environmental Protection Agency (EPA) released its external review draft of a document titled "Risk Assessment to Support the Review of the PM Primary National Ambient Air Quality Standards" (hereafter, the "Draft PMRA"). Following are my comments on issues in conducting a quantitative risk assessment for $PM_{2.5}$.

I. Introduction and Summary of Main Points

EPA's Draft PMRA is intended to provide quantitative estimates of the levels of risk from $PM_{2.5}$ actually being experienced by the U.S. population today, and how those risks will change if current ambient $PM_{2.5}$ is reduced by the application of more stringent National Ambient Air Quality Standards (NAAQS). The Draft PMRA only quantifies those risks that have been determined to be "causal" or "likely causal" in the second draft of the Integrated Science Assessment (hereafter, the "Draft ISA"). Once such a determination is made in the ISA, however, the Draft PMRA not only presumes that statistical estimates of $PM_{2.5}$'s relative risks are causal, but also that they can be interpreted quite literally as the quantitative concentration-response functions that determine actual risks. Whatever merit the observed epidemiological associations may have as indicators of a causal relationship, the unquestioning numerical credence that EPA assigns to the epidemiological estimates undermines the credibility and reliability of the Draft PMRA results. A common adage is "association does not imply causation;" to this can be added that even if a statistical association *does* reflect causation, it does not define the actual quantitative response function.

There are several layers of problems with the quantitative risk calculations in the Draft PMRA:

- At the most fundamental level, the Draft PMRA *presumes* that there is a causal association in the epidemiological evidence. That presumption is less than settled, as my comments will explain, because all of the studies may be wrong for the same systematic reason.

- At the next level, the estimates of relative risk from the epidemiological studies are almost certainly highly biased and, as my comments will explain, the bias is likely in the upward direction. Because the Draft PMRA uses those estimates directly for its quantifications of risk, the Draft PMRA's estimates of current premature mortality from $PM_{2.5}$ exposures are probably overstated.
- Topping it off, the epidemiological studies are incapable of defining how the relative risk would tend to vary at increasingly lower levels of $PM_{2.5}$. This creates an increasingly large error as risk reductions are estimated for tighter and tighter alternative ambient $PM_{2.5}$ standards.

All of the above problems in using the epidemiological relative risk estimates for quantitative purposes stem from fundamental data limitations that face every single one of the epidemiological research teams. This situation does not imply any fault on the part of the research teams or that the quality of their work is not of high quality. Unfortunately, however, EPA understates the remaining uncertainties that result from these studies' inherent data limitations when it engages in the quantitative risk assessment in its Draft PMRA.

Given that it does not address this array of problems in using epidemiological studies to attempt to quantify risk, the quantitative risk assessment in the Draft PMRA is highly misleading as an input to policy decisions. EPA could mitigate this situation by finding ways to quantitatively incorporate corrections for the systematic biases. This would produce a larger range of uncertainty in its estimates of risk, but one that reflects the true current state of knowledge. If this is not done, however, then the Draft PMRA should not be used in the consideration of alternative $PM_{2.5}$ NAAQS.

The rest of my comments are organized in the following way:

- Section II explains the problem in making a presumption of causality in the Draft PMRA, even though this is the determination in the Draft ISA.
- Section III discusses how the data limitations of the available $PM_{2.5}$ epidemiology literature make it inappropriate to use the estimated relative risks from those studies directly in a quantitative risk assessment as the Draft PMRA does.
- Section IV points out that even the epidemiological studies indicate a non-negligible chance that $PM_{2.5}$ imposes no long-term risk to all-cause mortality at all, once they are reviewed in a less selective manner than in the Draft PMRA.
- Section V summarizes and concludes that until these uncertainties are addressed quantitatively in the PMRA, it will be unreliable, and should not be used in the consideration of alternative $PM_{2.5}$ NAAQS.

II. EPA's Causality Criteria Are Inappropriate and Promote False Confidence in Quantified Risk Estimates in the Draft PMRA.

The Draft ISA provides a review of the weight of evidence in favor of alternative degrees of causal inference for various effects, such as between long-term exposures to $PM_{2.5}$ and cardiovascular mortality. The Draft ISA's determinations of causality for long-term $PM_{2.5}$ associations with CVD mortality and likely-causal for all-cause mortality associations are heavily driven by evidence of statistical associations in observational epidemiology studies. As explained below, this degree of reliance on the epidemiological evidence is excessive, highlighting a weakness in the criteria for causality that EPA establishes in that document. Given that uncertainties in the causality determination are never questioned again in the Draft PMRA, they are important to discuss in these comments on the Draft PMRA.

(II.A.) EPA's criteria allow excessive reliance on epidemiological findings in making a determination whether pollutants are causally associated with health effects.

The Draft ISA provides a set of criteria that must be met to determine that a particular health effect is causally related to $PM_{2.5}$ exposure. I quote the criteria below, which I have broken into two parts for purposes of the discussion that follows:[1]

Part 1: "The pollutant has been shown to result in health effects in studies in which chance, bias, and confounding could be ruled out with reasonable confidence. For example: a) controlled human exposure studies that demonstrate consistent effects; or b) observational studies that cannot be explained by plausible alternatives or are supported by other lines of evidence (*e.g.*, animal studies or mode of action information)."

Part 2: "Evidence includes replicated and consistent high-quality studies by multiple investigators."

On its own, Part 1 provides what would seem to be a perfectly appropriate set of criteria for making a causal determination. However, the addendum of Part 2 greatly weakens the requirements, because as a logical "or" statement, it provides a way for EPA to conclude in favor of causality even if controlled human exposures studies do *not* demonstrate consistent effects, *and* observational studies *can* be explained by plausible alternatives *and* are not supported by animal studies or mechanistic actions; Part 2 allows EPA to conclude in favor of causality even in the face of all of the foregoing findings as long as multiple authors have published quality epidemiological studies that replicate each other. Thus, the single sentence of Part 2, treated as a logical "or" rather than as a logical "and" to the requirements of Part 1, serves to absolve EPA from having to demonstrate that the associations in chronic studies "cannot be explained by plausible alternatives" before it can make its causal determination.

[1] Draft ISA, Table 1-3, p. 1-29.

Unfortunately, EPA relies almost entirely on the type of evidence allowed by Part 2 to make its causal determination regarding long-term mortality risks of $PM_{2.5}$. Take the case of long-term cardiovascular (CVD) mortality risk, for example. The Draft ISA relies almost entirely on the existence of a multiplicity of separate chronic studies that all find a $PM_{2.5}$-mortality association. It begins and ends that causality discussion with the following respective quotes:

> "A number of large, multicity U.S. studies (the ACS, Six Cities Study, WHI, and AHSMOG) provide consistent evidence of an effect between long-term exposure to $PM_{2.5}$ and cardiovascular mortality."[2]

and:

> "In summary, a number of large U.S. cohort studies report associations of long-term $PM_{2.5}$ concentration with cardiovascular mortality. These studies provide the strongest evidence for an effect of long-term $PM_{2.5}$ exposure on CVD effects."[3]

EPA determines that this evidence from multiple chronic studies suffices to identify a causal relationship, even though the remainder of the supporting evidence that the Draft ISA musters in support of a long-term CVD mortality risk are some inconsistent morbidity studies, inconsistent clinical studies, and a few toxicological studies that are suggestive of some possible relevant responses. The Draft ISA does not offer any reasons to believe that any of the many observational studies have met EPA's conditions for making a causal determination under Part 1. Nowhere does EPA make a case that the associations in the chronic mortality studies "cannot be explained by plausible alternatives." In fact, it could not possibly make such a claim given that epidemiologists continue actively to try to rule out various plausible alternatives. The plausible alternatives that researchers have not yet been able to rule out include confounding (residual or otherwise) by co-pollutants, noise, stress, and socioeconomic factors.[4] For example, the most recent ACS cohort analysis (Krewski *et al.*, 2009) focused vigorously on more effectively controlling for socioeconomic factors than in past studies, but it made no attempt to rule out possible confounding by co-pollutants.

Thus, the available evidence for long-term mortality risk does not meet the causality criteria contained within Part 1, and EPA is relying very heavily on Part 2 to defend its "causal" determination. Part 2, however, is not a valid basis for a causal determination if observational studies are not also supported by consistent effects in controlled studies. It would be a reasonable addendum if it were required *in addition* to Part 1 conditions, but not when used *instead of* meeting Part 1 conditions. The reason it is insufficient for

---

[2] Draft ISA, p. 7-25.
[3] Draft ISA, p. 7-26.
[4] The effect modification by educational status in the chronic studies remains as an indication of some residual confounding by socioeconomic factors. Although this pattern was reduced in the Krewski *et al* (2009) study, it was not eliminated.

identifying whether a statistical association is causal is because it fails to address the possibility of *systematic biases*, which cannot be ruled out except with evidence from controlled studies.

Systematic biases will occur if the studies in question have relied on similar methodologies and similar data sources. In this case, if a bias (*e.g.* due to residual confounding) exists in one study, then it is likely to exist in all the studies. All of the epidemiological results can be wrong for the same reason. A multiplicity of studies finding a statistical association do not provide independent confirmation supporting a causal inference unless one can demonstrate that there is no potential for such systematic bias among those multiple studies. Thus, Part 2 of EPA's causality criteria enables causality to be declared even if there remains a very large likelihood of no causality.

There is substantial potential for systematic bias in the case of chronic $PM_{2.5}$-mortality epidemiology.[5] Table 7-8 of the Draft ISA lists 14 recent U.S. cohort studies that find an association between $PM_{2.5}$ and mortality.[6] All of these studies draw from the same fundamental data set, however, because they all sample individuals across the U.S. and assess the correlation between their local monitors' $PM_{2.5}$ levels and their mortality risks after attempting to control as best possible for the very broad swath of much stronger determinants of risk (*e.g.*, age, sex, diet, smoking habits, and socioeconomic factors). Controlling effectively for these other factors is the key to getting a sound answer, y*et all* of the studies are reduced to using approximately the same approximate data, all of them facing enormous amounts of error in how those variables are assigned to individual cohort members. In any single study, there is a good chance that the controls for the primary determinants of mortality risk are incomplete, and some confounding remains to bias the association estimated for $PM_{2.5}$. Unfortunately, all of these studies face the same problem, in a systematic way, because they all rely on the same types of data, and face the same fundamental data limitations.[7]

The fact that these studies rely on several different cohorts does not make them independent of each other with respect to confounding and effect modification. If ambient $PM_{2.5}$ is correlated with the missing or poorly measured non-PM explanatory variables across the U.S., then almost any reasonably diverse subset of the U.S population are likely to embody that same underlying correlation. For example, a sample of mostly white volunteers for a cancer study and a sample of veterans may have quite different socioeconomic profiles, but both will reflect the general correlations that exist across the U.S. between $PM_{2.5}$ measured at central monitors and key socioeconomic or other non-$PM_{2.5}$ causal factors. The same biases can apply to every single cohort of people drawn from the U.S. population. This is a particular systematic concern for

[5] Goodman (2009), pp. 8-9, makes a similar point.
[6] Draft ISA, pp. 7-119 to 120.
[7] There is also uniformity in the methodologies being used, in that almost all of the researchers use the same statistical model, the Cox proportional-hazards (PH) model, and thus systematically share any biases that may derive from the limitations of this statistical model. For example, the Cox PH model assumes the effects of pollution levels and of potential confounders on the logarithm of hazard are all linear. The assumption of proportional hazards has received limited testing, but that which has been done raises serious questions about this key assumption (see, for example, Abrahamowicz *et al.*, 2003).

variables that must be controlled at the ecological level. Thus, the biases from confounding and effect modification that are most difficult to control are also likely to be systematic across multiple U.S. cohorts.

This is not a criticism of the quality of the research teams' efforts; it is just an unfortunate reality of the limitations of the available data and tools to study such a subtle possible risk without the ability to perform controlled experiments. Nevertheless, the potential for systematic bias should not be ignored.

(II.B.) Differential measurement error can systematically create bias from confounding, even if confounders are included in the epidemiological study.

The potential for systematic confounding and effect modification cannot be eliminated simply by including the relevant explanatory variable in the data analysis if there is measurement error. A confounder is a variable that has a direct effect of its own on risk, *and* which is correlated with the pollutant being studied. When this is the case, if the confounder is left out of the analysis, the pollutant will be attributed some of the explanatory power that is actually due to the missing confounder, and thus the relative risk estimate will be biased. If personal exposures to all the confounders and effect modifiers can be measured accurately, the bias in the $PM_{2.5}$ exposure-risk association can be eliminated by identifying and including the confounder in the data analysis. However, if the confounder cannot be measured with accuracy, even if it is included in the data analysis, *residual* confounding will remain in the estimated association between the exposure variable and risk.

The situation is complex, but simulation studies can help understand the potential for biased effects estimates in situations that have confounders with differential measurement error. One such study (Fewell *et al.*, 2007) demonstrates that typical amounts of confounding, combined with typical amounts of measurement error, can cause quite large relative risks to be assigned to an exposure variable that has no effect at all, even when measures of the confounder are included among the controlling covariates in the analysis. The size of the potential erroneous relative risk reported by Fewell *et al.* exceeds the magnitude of the $PM_{2.5}$ relative risk in the $PM_{2.5}$ epidemiological literature. The implication is that the estimated chronic $PM_{2.5}$ relative risks – given their fairly small magnitude – could be the result of residual and unmeasured confounding by either socioeconomic factors or other environmental factors. Because these potential confounding relationships are likely to exist for every cohort if they exist for one, all of the multiple, independent long-term $PM_{2.5}$-risk associations could be reflecting the same systematic biases (Boffeta *et al.*, 2008).

Cohort studies in other countries might not face all of the same systematic errors that would apply to cohort studies all from the U.S. However, if the source of the confounder is physically related to PM sources, then one would see the same systematic bias even for cohort studies outside of the U.S. Thus, positive findings from cohort studies in other countries might reduce some of the concern with a socioeconomic-pollutant correlation, but cannot eliminate concerns that PM is a proxy for another co-emitted pollutant, some

non-chemical effect coinciding with chemical emissions (*e.g.*, noise), or for a single constituent of PM. Table 7-8 of the Draft ISA identifies one $PM_{2.5}$-mortality cohort study from the Netherlands (Brunekreef *et al.*, 2009). The abstract for this study states that it "differs from cohort studies based on city-level differences in exposure" because it considered exposures to pollution sources that exist mainly *within* cities. The U.S. cohort studies, which are based on city-level differences in pollution, are reporting an association between health and the total, undifferentiated mass of *all* forms of $PM_{2.5}$. It is those total mass associations that are then being used for risk calculations in the Draft PMRA. Thus, this non-U.S. study does not provide corroborating evidence of the same kinds of relationships being estimated by the U.S. cohort studies, and so it does not help eliminate concerns that potential systematic biases may pervade the latter.

(II.C.) Other pollutants are a likely source of residual and unmeasured confounding bias in the long-term associations.

The potential for the long-term $PM_{2.5}$ estimates to all share bias due to confounding from other pollutants is clear, regardless of what one might think about residual socioeconomic variable confounding in this body of literature. Rarely are other pollutants or pollutant-sources included in $PM_{2.5}$ epidemiological regressions; however, when they are included, there is often a marked reduction in the size and statistical significance of the $PM_{2.5}$ effect. Such a sensitivity upon the inclusion of SO_2 was a major finding of the reanalysis of the ACS data by Krewski *et al.* (2000); yet, the subsequent papers of Pope *et al.* (2002) and Krewski *et al.* (2009) that extended the ACS cohort analysis did not report any $PM_{2.5}$ relative risks that had also been controlled for SO_2. This omission in recent ACS-based studies leaves an important question regarding the quantitative validity of the $PM_{2.5}$ associations taken from Krewski *et al.* (2009) that the Draft PMRA uses.

The other study that EPA places high reliance on is the update of the Harvard Six Cities study by Laden *et al.* (2006). EPA cites this study as providing confirmatory evidence that long-term reductions in $PM_{2.5}$ produce a corresponding reduction in mortality. However, this study does not consider any pollutants other than $PM_{2.5}$, even though the levels of various gaseous pollutants have fallen concomitantly with $PM_{2.5}$ in the six cities. Rather, Laden *et al.* attribute the entire change in health risk associated with air pollution reduction to $PM_{2.5}$ without any apparent attempt to test whether any other pollutant might have equivalent explanatory value.[8] The paucity of data points available with this cohort make it impossible for statisticians to even attempt to control for more than one pollutant at a time. That is, with the Harvard Six Cities cohort, the estimate of the effect of pollution on inter-city mortality risk differences must be based on only 6 cities/data points. In contrast, as many as 150 cities/data points are available for inferring relative risk estimates with the ACS data set. Nevertheless, the researchers still could have used a series of one-pollutant models to explore whether pollutants other than $PM_{2.5}$ might also be associated with the observed changes in inter-city mortality risks.

[8] Furthermore, in the critical second period of this study, the results were not based on actual fine PM measurements. Rather, they are rather based on measurement of another NAAQS pollutant (PM_{10}) and extinction coefficients.

Even if other pollutants were to be included in these long-term risk studies, it is quite likely that they would fail to control for confounding because of differential measurement error. $PM_{2.5}$ is generally believed to have a much more uniform distribution in space than other pollutants with which it is correlated, including NO_2, CO, coarse PM, and even ozone. Thus, the standard practice of using data from a central monitor to estimate individuals' exposures probably results in greater exposure misclassification for these other pollutants than for $PM_{2.5}$. The result could be that $PM_{2.5}$ will persistently appear to carry the best explanatory power, yet just be serving as a proxy for the health effects of the other, more erroneously measured pollutant exposures. This possibility was demonstrated in a simulation study that contained considered two hypothetical correlated pollutants, one a "True Culprit" measured with relatively large error, and the other an "Innocent", but measured with relatively small error. The simulation results showed that:

> "in circumstances like this, *which* pollutant would appear to have the most significant and consistent relationship with health may be determined more by its relative observation error than by its actual contribution to the health effects in question. The greatest problem with this spurious association of Innocent with health is that it remains stable whether or not True Culprit is added into the regression. Further, if only True Culprit is included in the regression, the R^2 falls to zero. True Culprit is the pollutant that seems to have a highly unstable association and very little explanatory power on its own. Thus, unlike in the simple confounding case, the usual methods for checking for confounding no longer function well when there are observation errors as well as strong correlation among pollutants."[9]

(II.D.) Evidence of proxy effect exists in changes in estimated $PM_{2.5}$ relative risk coefficients over time.

One signal that a non-causal proxy effect might account for the $PM_{2.5}$-mortality associations would occur if the estimated relative risk for $PM_{2.5}$ mass were to increase over time as $PM_{2.5}$ levels decline. That is, if there is a given amount of risk associated with a certain non-$PM_{2.5}$ causal factor that is correlated with $PM_{2.5}$ mass, if the unidentified causal factor was not reduced while $PM_{2.5}$ was reduced, then the remaining lower levels of $PM_{2.5}$ would account for the same total level of risk from the unidentified factor. The result would be a greater relative risk associated with a given amount of PM difference.[10]

---

[9] Smith and Chan (1997), p. 21. (The observed correlation between the two pollutants in this analysis was 0.56, which is in the range often observed in the US.)

[10] This signal will not necessarily occur even if $PM_{2.5}$ is serving solely as a proxy for some unnamed causal factor, if the causal factor were to be reduced in roughly the same degree as $PM_{2.5}$ over that same period of time. That could be the case if $PM_{2.5}$ is serving as proxy for a gaseous pollutant, since most of the pollutants have been declining simultaneously due to parallel environmental regulations.

We do observe this pattern in the extended analyses of the ACS cohort. For example, in Pope *et al.* (2002), the estimated relative risk per 10 $\mu g/m^3$ of $PM_{2.5}$ for all-cause mortality rose from 1.04 when using 1979-1983 $PM_{2.5}$ data (which averaged 21.2 $\mu g/m^3$) to 1.06 when using 1999-2000 data (which averaged 14.0 $\mu g/m^3$). The estimated relative risk for cardiopulmonary risks rose from 1.06 to 1.08.[11] Similarly, in Krewski *et al.* (2009), the all-cause relative risk rises from 1.043 to 1.056 and the cardiopulmonary relative risk rises from 1.089 to 1.129 when estimated with the earlier or later $PM_{2.5}$ data.[12] (Laden *et al.* (2006) report a decline in the relative risks from an earlier period to a later period of exposure, but this is not the same comparison. In the ACS examples, the risk over the same follow up period is estimated using $PM_{2.5}$ from two different parts of the time period. Laden *et al.* do not report estimates or relative risk for the entire time period using first the earlier, then the later $PM_{2.5}$ measures. Since they do not report a comparable set of relative risks, their finding cannot be said to conflict with the ACS finding just described.)

The increase in the estimated relative risk that occurs in the ACS data set when more recent $PM_{2.5}$ data are used might also occur if there is a real effect from $PM_{2.5}$ mass that is truly long-term in nature. In that case, on-going mortality outcomes might be a function of earlier exposures to $PM_{2.5}$, when it was at higher levels, while more recent $PM_{2.5}$ measures might be serving as a proxy for the historically higher $PM_{2.5}$ exposures. Even if this does explain the upward trend in estimates of $PM_{2.5}$ relative risks, it implies that any quantitative estimate of benefits from reducing current $PM_{2.5}$ based on the numerical results of recent epidemiological associations will be biased upwards. That is, adoption of a relative risk estimated using the more recent $PM_{2.5}$ "at face value" as the quantitative $PM_{2.5}$ concentration-response function would be erroneously assuming that the entire increase in risk due to higher historical $PM_{2.5}$ exposure is caused by a much smaller amount of $PM_{2.5}$ exposure. This concentration-response function would overstate the risk from as-is $PM_{2.5}$, and it would overstate reductions in risk that could be expected by reducing today's lower $PM_{2.5}$ to yet lower levels.

In summary, proxy effects can be at play in chronic studies, even if there is a causal relationship for $PM_{2.5}$ mass generally, and this proxy effect would create erroneous (overstated) risk and risk-reduction estimates in the PMRA. The bias would then be exacerbated even further when considering the benefits of further reductions in $PM_{2.5}$ due to rollbacks to alternative, more stringent $PM_{2.5}$ standards. The latter possibility is discussed further in Section III.B.

(II.E.) Epidemiological findings on short-term mortality are far more heterogeneous, and do not provide strong back-up to long-term studies.

Some people prefer to rely on short-term, time-series studies for evidence of an effect from $PM_{2.5}$ because effects observed within each city provide more inherent control for

[11] Pope *et al.* (2002), Table 2, p. 1136.
[12] Krewski *et al.* (2009), Table 6, p. 23. Values reported are for regressions with MSA & DIFF ecological controls, but the pattern also appears in other regression in the table.

socioeconomic factors that are otherwise difficult to measure accurately. While this may be true to some extent for the socioeconomic factors, short-term studies are still subject to potentially uncontrollable confounding from other pollutants. Nevertheless, existing short-term risk studies offer very little support for a causal interpretation of the observed long-term $PM_{2.5}$ associations. In particular, the quantitative level of the risk in the long-term studies is roughly an order of magnitude higher than associations found in short-term studies. The difference could be entirely due to confounding bias, or – as EPA prefers to explain it – the difference could be that cumulative, long-term effects are much larger than acute effects. Neither explanation can be held up as more correct, but even EPA's preferred explanation implies that the short-term studies cannot be viewed as corroborating the long-term study findings, because it implies that the long-term associations would have to be for an effect that acute studies cannot even detect.

Short-term studies also produce results that vary enormously from city to city and regionally, often finding no effect at all, even in cities and regions with relatively high $PM_{2.5}$ levels. This heterogeneity may indicate that the smaller, short-term $PM_{2.5}$ associations are not necessarily causal either.

III. Even If the $PM_{2.5}$ Association Is Causal, Statistical Estimates of $PM_{2.5}$ Relative Risk Remain Subject to Biases that Make Them Unreliable for Quantifying Risks.

(III.A.) Biases in estimates of the average magnitude of the $PM_{2.5}$ association are likely due to four types of data problems.

As explained in Section II, the Draft ISA's conclusion that $PM_{2.5}$ is causally related to cardiovascular mortality risks (and likely causally related to mortality risks in general) remains open to reasoned debate. However, a variety of uncertainties also exist that directly undermine the *quantitative* interpretation of the epidemiological findings for determining what numbers of deaths are premature at present, and especially for predicting how mortality risks would change if $PM_{2.5}$ mass were reduced. There are at least four ways in which quantitative biases can be present in the epidemiologically-estimated associations that would undercut their reliability for quantification of risks, as discussed below.

(1) Differences in potency of various $PM_{2.5}$ constituents. There are uncertainties about the relative potency of different constituents within the $PM_{2.5}$ mass.[13] Thus, even if a relative risk estimate is quantitatively valid as an *average* effect of the current mix of $PM_{2.5}$, if some constituents would not be reduced as much as others when an alternative $PM_{2.5}$ standard is imposed, then the reduction in risk from that standard would not be what one would predict using the *average* relative risk. In fact, if some small subset of the mass is highly potent and accounts for most or all of the observed association, it is

[13] The Draft ISA states (at p. 7-129) that "only a very limited number of the chronic exposure cohort studies have included direct measurements of chemical-specific PM constituents other than sulfates, or assessments of source-oriented effects, [in] their analyses." Also (at p. 2-25): "It remains a challenge to determine relationships between specific constituents, combination of constituents, or sources of $PM_{2.5}$ and the various health effects observed."

quite likely that this culprit would escape implementation plans, which will naturally be focused on reducing the constituents that account for the largest portions of the mass. The result could be no risk reduction at all, despite reductions in $PM_{2.5}$ mass; yet the Draft PMRA's methodology would predict substantial benefits from the tighter standard.

(2) Missing or inaccurate socio-economic variables correlated with regional $PM_{2.5}$ levels. All of the epidemiological studies have taken steps to provide controls for socioeconomic variables that affect mortality risks, but information is insufficient to ensure that these have been fully specified; also, many of the potential socioeconomic confounders and effect modifiers can only be measured with substantial amounts of error. Sometimes these errors can be found and eliminated through careful data quality work.[14] However, the errors of concern here are for socioeconomic data that are simply not possible to obtain. For example, although various updates of the key ACS study have been published up through 2009, they rely on the same individual and socioeconomic data collected in 1982, almost 30 years ago. Thus, it is not possible to assess changes in key confounding factors such as smoking cessation rates that are well known to fall along socioeconomic lines.[15] Other socioeconomic variables, such as data on the degree of stress in family life, are simply not possible to obtain and will never be possible to control for in the chronic risk studies.

Thus, despite extensive socioeconomic controls in all of the chronic risk studies, there remains the possibility that $PM_{2.5}$ mass is at least partially serving as a proxy for unidentified, or poorly measured, socioeconomic variables. If so, the $PM_{2.5}$ risk coefficient is biased. In this case, while it is not certain what the direction of bias would be, it is likely to be in the upward direction because lower socioeconomic status tends to be positively correlated with mortality risk and also with living in areas with higher pollution.[16] Regardless of direction of bias, quantified risk estimates that use the estimated $PM_{2.5}$ relative risk "at face value" will be incorrect.

(3) Other pollutants, even if included in the analysis. Much has been said about the possibility that $PM_{2.5}$ is serving as a proxy for another pollutant that has the true causal role. Studies have, at times, considered the role of other pollutants but this practice has been inconsistent.[17] When multi-pollutant results are not reported, one never knows if

[14] For example, the reanalysis by Krewski *et al.* (2000) of the Harvard Six Cities Study found that the coding protocol allowed cigar and pipe smokers to be classified as "non smokers;" the calculation of pack-years of smoking cigarettes was inconsistent, resulting in an underestimate of smoking pack-years of about 3% in some cities; and the error rate for the education variable on the earliest form used was 18%; etc. These kinds of errors can be avoided through careful review, and fixed, if detected.

[15] Some hypothesize that the errors in data on smoking cessation might explain the education gradient in $PM_{2.5}$ mortality observed in this study as well as the Harvard Six Cities Study; if so, the $PM_{2.5}$ relative risk estimate is probably a biased estimate.

[16] In the one known example where those with higher socioeconomic status happen to live in an area where the $PM_{2.5}$ is higher (*i.e.*, New York City), the lack of any increased mortality risk attributed to exposure to $PM_{2.5}$ in this group, versus those with lower socioeconomic status and lower $PM_{2.5}$ exposure, may illustrate the importance of socioeconomic confounding in air pollution epidemiology studies (Krewski *et al.* 2009).

[17] The Draft ISA states (at p. 7-82): "Given similar sources for multiple pollutants (e.g., traffic), disentangling the health responses of co-pollutants is a challenge in the study of ambient air pollution."

they may have been performed and found to have had the effect of attenuating the reported association for $PM_{2.5}$. Even if those studies simply did not perform any multi-pollutant regressions, one must wonder, why not? For example, the most recent papers based on the ACS cohort (Pope *et al.*, 2002; Krewski *et al.*, 2009) did not attempt to explore whether SO_2 had greater explanatory power than $PM_{2.5}$ mass, even though this was a widely discussed source of sensitivity reported in the preceding ACS paper (Krewski *et al.*, 2000).[18] Thus, the most recent epidemiological studies are not necessarily the most thorough in their efforts to explore confounding by other co-pollutants; as a result, their quantitative estimates cannot be viewed as more reliable for use in a quantitative assessment of $PM_{2.5}$-related risks. Relative risk estimates from the earlier studies that *did* account for co-pollutants could be less biased than the more recent relative risk estimates, even though the earlier ones have less statistical power due to their reliance on shorter cohort follow-up periods.

(4) Other environmental factors. Gaseous pollutants are not necessarily the only other non-$PM_{2.5}$ environmental factor for which $PM_{2.5}$ might be serving as a proxy. Measures such as proximity to traffic and intensity of local traffic have been the subject of much recent exploration of the basis for the $PM_{2.5}$ associations (*e.g.*, Lipfert, Wyzga *et al.*, 2006; Lipfert, Baty, *et al.*, 2006; Jerrett *et al.*, 2005; Beelen *et al.*, 2008a). In most cases where they have been accounted for in the data analysis, traffic-related variables appear to have the stronger associations. This could point to certain $PM_{2.5}$ constituents, or to some gaseous pollutants, and it could point to other factors such as noise or stress (Lipfert, Baty *et al.*, 2006; Beelen *et al.* 2008b).[19] Rolling back $PM_{2.5}$ mass would not have any effect on these other possible causal factors; again, the quantitative interpretation of a $PM_{2.5}$ relative risk would produce completely erroneous estimates of risk reductions from alternative $PM_{2.5}$ standards.

The point is obvious: even if the associations with $PM_{2.5}$ have a causal element, the many limitations of the epidemiological data mean that the relative risks estimated by epidemiological studies do not offer a direct quantitative relationship for how $PM_{2.5}$ mass alone affects either current health risks, or changes in risks if $PM_{2.5}$ mass is reduced. In all of the situations described above, the observed association could be "statistically significant," yet the health benefits from rolling back $PM_{2.5}$ could be as low as zero, because statistical significance calculations cannot detect the presence of bias; in fact, they presume it does not exist. Also, if differential measurement errors are at work, then

[18] HEI's commenters on the 2009 Krewski *et al.* study lament the lack of further study of confounding by copollutants, but offer their own excuse for this omission: "Given that the Reanalysis (Krewski *et al.*, 2000) had extensively tested the potential for the gaseous pollutants to confound the relationship between exposure to $PM_{2.5}$ and mortality and had not found any significant confounding (other than by SO_2), it is understandable that the current investigators chose to focus their limited resources on the extensive exploration of spatial autocorrelation in a series of one-pollutant models." (Krewski *et al.*, 2009, p. 130, emphasis added). This is a rather weak reason if their goal is to explore the strength of the $PM_{2.5}$ mass association in greater detail, given that their previous paper's findings on that association were the most sensitive to the inclusion of SO_2 as a co-pollutant.

[19] Bukowski (2008) has suggested that noise and stress could be an uncontrolled factor also affecting short-term $PM_{2.5}$-exposure risk studies.

one cannot have confidence that the bias would be eliminated, or even mitigated, by having included these other factors in the epidemiological regressions.

(III.B.) The magnitude of the $PM_{2.5}$ association at varying $PM_{2.5}$ exposure levels (i.e., the "shape" of the relationship) is estimated with error.

Brauer *et al.* (2002), among others, have demonstrated that measurement error for personal exposures when using central monitor data can hide a threshold from the statistical methods, even when one exists in reality.[20] Thus EPA's assumption of a linear relationship is not valid for quantitative risk analysis, even if the estimated constant relative risk were quantitatively valid as an average over the range of observed exposure levels.

EPA does not consider concerns with measurement error at all when it concludes that a linear relationship "most adequately" represents the association for purposes of statistical fit. When the Draft ISA states that "the C-R curve was found to be indistinguishable from linear,"[21] it is only making a statement about statistical goodness of fit to the available data. These statistical tests offer no information about whether the shape of the underlying true relationship has been obscured by measurement error, yet this is the critical question when trying to develop a quantitative estimate of risk from statistical associations using messy data.[22] The risk quantification step of the PMRA requires a functional relationship that properly reflects the true shape of the underlying relationship in order to reliably predict the changes in risk that would result from changes in $PM_{2.5}$. If a risk relationship has a non-linear shape (as one would logically expect for a true concentration-response situation, given a normally distributed degree of sensitivity across a population), then it is not appropriate to simply use a linear relationship just because the available epidemiological data do not offer the sensitivity necessary to detect that shape.

Errors due to an incorrect shape of the concentration-response function will become more and more pronounced with rollbacks to increasingly lower levels of $PM_{2.5}$ because the amount of change in risk associated with increasingly lower $PM_{2.5}$ exposures may

---

[20] Smith and Chan (1997) also demonstrated the impossibility of statistically detecting a real threshold in the presence of exposure measurement error. For simulated data with a pronounced ("hockey-stick") threshold at 20, the best fit for alternative thresholds was no threshold at all. They also fit a nonparametrically smoothed curve to the simulated data, with the resulting estimated relationship appearing to have the opposite of a threshold, that is, that the estimated concentration-response curve became *steeper* at concentrations closer to zero (and well below the point of the actual threshold). See Figure 9 and associated discussion in Smith and Chan (1997), p. 14, for the nonparametrically smoothed estimate of the concentration-response curve.

[21] Draft ISA, p. 2-37.

[22] Even the conclusion that linearity is the best *statistical fit* remains a debatable conclusion: see, for example, Abrahamowicz *et al.* (2003), and Goodman (2009), pp. 21-22. Without even considering debates about statistical tests of nonlinearity, potential evidence of nonlinearity can be found in the extended cohort analyses. For example, the finding reported by Laden *et al.* (2006) that estimated relative risks were lowered as $PM_{2.5}$ levels fell over time implies a non-linear relationship, while Gamble and Nicolich (2006) argue that a non-linear relationship may be observable even in the relative risks for a single time period. Also, the ACS evidence discussed in II.D that recent $PM_{2.5}$ levels used to estimate long-term associations could be serving as a proxy for earlier, higher $PM_{2.5}$ exposures also implies a non-linear actual relationship.

become vanishingly small, while the presumed linear statistical association assumes equally large amounts of risk reduction for a unit of improvement in $PM_{2.5}$, whether that change occurs from high levels of as-is exposure, or from near-background levels. The practice in the Draft PMRA of not counting risks below the lowest measured level (LML) of $PM_{2.5}$ does not eliminate this quantitative error. In fact, this practice exacerbates the misleading nature of the PMRA because it produces the bizarre effect of suggesting that there is no threshold, *yet also* that 100% of the currently existing risk attributed to $PM_{2.5}$ would be eliminated if $PM_{2.5}$ is rolled-back only as far as the LML.

(Setting aside the logical inconsistency, if one truly believes that a linear relationship can be assumed, then the Draft PMRA is misleading when it reports that nearly 100% the long-term $PM_{2.5}$ risk can be eliminated by tightening the standard for $PM_{2.5}$ to 12 $\mu g/m^3$ annual average. All EPA can really say is that the latter standard would reduce $PM_{2.5}$ to the lowest levels that were observed in the most recent years among the cities included in the ACS cohort studies.)

IV. The Epidemiological Evidence Itself Indicates a Meaningful Chance that Long-Term $PM_{2.5}$ Exposures Do Not Elevate Public Health Risk.

Sections II and III have provided multiple reasons why quantitative estimates of risk in the Draft PMRA are unreliable. The Draft PMRA uses estimates of relative risks from the epidemiological studies at "face value" for its quantitative concentration-response functions, ignoring the many likely sources of bias those estimates. At a minimum, this results in large errors in its estimates of risk and risk reductions from reducing ambient $PM_{2.5}$ exposures that are not captured in statistical confidence intervals. However, EPA also commits an error of omission in its Draft PMRA by relying selectively on a few relative risk estimates from a few studies. This hides the degree of uncertainty that is detectable in the full body of epidemiological evidence, even if taken "at face value."

Even if one accepts the existing body of long-term epidemiological relative risk estimates at face value, one can observe a substantial chance that no effect exists at all if the full set of available relative risk estimates is given fair consideration. Figure 7-6 of the Draft ISA (p. 7-123) shows a very selective set of results. However, the entire body of evidence includes past as well as current studies. Even earlier estimates from the ACS studies remain relevant to the extent that certain regressions in them have not been reproduced in more recent ACS analyses. (The results that include SO_2 as well as $PM_{2.5}$ in the regressions from Krewski *et al.*, 2000, are a case in point.) Other studies also not shown in Figure 7-6 have produced non-positive and/or no statistically significant association between $PM_{2.5}$ and chronic mortality.[23] Nothing in the more recent literature

[23] Lipfert *et al.* (2000) is an example that found a negative association. More recent papers for the Veterans cohort (*e.g.*, Lipfert, Wyzga, *et al.*, 2006) have reported positive $PM_{2.5}$-risk associations (although significant only in 1-P formulations) but these newer findings do not make the earlier findings irrelevant. The earlier findings used a different subset of the Veteran's Cohort than the later findings, where the subset was driven by the locations of the available pollution data being used. The earlier studies also considered mortality risk in a different (earlier) time period. Thus, one can find both negative and positive associations within this single cohort, depending solely on the time period and air pollution data used.

necessarily supersedes those other studies. When the entire set of regressions are considered, giving equal weights to single and multipollutant models, and giving equal weights to the various cohorts that have been studied, one finds that this literature *as a whole* suggests that there is about a 15-20% chance that there is zero risk from $PM_{2.5}$.[24] This reflects the fact that the literature does contain quite a few findings of insignificance, which imply, based on statistical error alone, that the effect is zero. (In fact, it implies a possibility of a negative effect, but for this discussion, those are considered zero, not beneficial, effects.)

V. Conclusion

EPA's criteria for determining whether a causal relationship exists between exposures to $PM_{2.5}$ and health endpoints are fundamentally flawed because they allow an incorrect determination of causality to be made in circumstances that are marked by systematic biases. The epidemiological literature for long-term $PM_{2.5}$ exposure risks is clearly open to the possibility, and even likelihood, of systematic bias. Therefore, EPA's "causal" determination for long-term cardiovascular mortality risk and its "likely-causal" determination for long-term all-cause mortality risk in the Draft ISA are subject to a much greater chance of being wrong than the words themselves suggest. This fact alone places significant limitations on the usefulness and reliability of EPA's quantitative estimates of long-term mortality risks in its Draft PMRA, because the entire risk assessment is predicated on an unquestioned *presumption* of causality.

The quantitative estimates of mortality risk in the Draft PMRA remain unreliable, however, even if it is correct that there is some causal relationship between $PM_{2.5}$ and risk of dying. This is because the Draft PMRA also presumes that the statistical associations in the epidemiological literature can be interpreted *literally* as the actual concentration-response relationships for quantifying current levels of risk, and changes in risks for altered ambient $PM_{2.5}$ conditions.

The translation from an epidemiologically-derived association to a real "concentration-response function" that quantifies how much risk would change if PM mass were changed is highly problematic, regardless of the quality of the epidemiological studies that are being relied on. Even if one has great confidence that an association between $PM_{2.5}$ and health risk detected in an epidemiological study is reflecting a true causal relationship, the statistical model and its parameter estimates (*e.g.*, the "relative risk") cannot be assumed to be a precise numerical estimate of the true causal relationship, given the many limitations of the available data. As explained above, there remain good reasons to suspect that some or all of the estimated association bears no causal implication for $PM_{2.5}$ mass itself, or even of one of its constituents. If there are any missing explanatory variables – which is almost certainly true – then the statistical estimate of relative risk is not quantitatively reliable to assess either "as-is" risks, or

[24] This includes consideration of relative risks in Eftim *et al.* (2008), Enstrom (2005), Jerrett *et al.* (2005), Krewski *et al.* (2000, 2009), Laden *et al.* (2006) Lipfert *et al.* (2002), Lipfert, Baty *et al.* (2006), Lipfert, Wyzga *et al.* (2006), McDonnell *et al.* (2000), Pope *et al.* (2002), and Villeneuve *et al.* (2002).

changes in risk as $PM_{2.5}$ is reduced. The kinds of measurement errors and confounding that are present in the $PM_{2.5}$ epidemiological data also mean that the shape of the true relationship cannot be identified. The inability to define the shape of the true relationship means that one can have no confidence in statements of how risk will change as $PM_{2.5}$ is reduced.

The statistical confidence intervals that the Draft PMRA offers up as "uncertainty" do not measure biases due to missing variables or measurement error, and thus do not offer a way of characterizing the numerical uncertainty in actual risk levels at current or rolled-back $PM_{2.5}$ levels. The Draft PMRA's sensitivity analyses, which simply substitute one statistical estimate of relative risk for another, also cannot begin to characterize the quantitative uncertainty, given that all the available epidemiological estimates suffer from the same limitations in data and methods, and are thus subject to a systematic bias. As a result, the numerical estimates provided in the Draft PMRA have no reliable relationship to reality, even if one accepts the presumption that the epidemiological studies are detecting a causal relationship between one or more constituents of $PM_{2.5}$ and health risk. With all of these unstated and unanalyzed presumptions, one can have no confidence in the estimates of how much health risk is being created by current $PM_{2.5}$ exposures, nor whether any of that estimated risk would be reduced by reducing an undifferentiated measure of total $PM_{2.5}$ mass. The Draft PMRA's quantitative estimates of risk from as-is $PM_{2.5}$, and quantitative estimates of reductions in due to lowered $PM_{2.5}$, are unreliable.

Until it contains a more explicit analysis of the quantitative implications of these inherent challenges for estimating risks under current and alternative ambient $PM_{2.5}$ standards, the quantitative risk assessment of the Draft PMRA is, at best, not useful; at worst, its results are highly misleading as an input to policy decisions for setting a NAAQS. The only way to obtain reliable estimates would be to find ways to quantitatively incorporate corrections for systematic biases due to differential measurement errors, and potentially unmeasured causal confounders of a non-pollutant nature. This would produce a larger range of uncertainty, but one that reflects the true current state of knowledge. If this is not done, however, then the Draft PMRA should not be used in the consideration of alternative $PM_{2.5}$ NAAQS.

References

Abrahamowicz M, Schopflocher T, Leffondre K, du Berger R, and 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. Toxicology and Environmental Health*, Part A, 66:1625-1653.

Beelen R, Hoek G, van den Brand P, Goldbohm R, Fischer P, Schouten L, Jerrett M, Hughes E, Armstrong B, Brunekreef B. 2008a. "Long-term effects of traffic-related air pollution on mortality in a Dutch Cohort (NLCS-Air Study)," *Environmental Health Perspectives* 166(2):196-202.

Beelen R, Hoek G, Houthuijs D, van den Brandt PA, Goldbohm RA, Fischer P, Schouten LJ, Armstrong B, Brunekreef B. 2008b. "The Joint Association of Air Pollution and Noise from Road Traffic with Cardiovascular Mortality in a Cohort Study," *Occup. Environ. Med.* (published online, November 18).

Boffeta P, McLaughlin, JK, LaVecchia C, Tarone RE, Lipworth L, Blot WJ. 2008. "False Positive Results in Cancer Epidemiology: A Plea for Epistemological Modesty," *J. of the National Cancer Institute* 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 Analysis* 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,* Research Report #139, Health Effects Institute, Boston, Massachusetts.

Bukowski J. 2008. "Do Pollution Time-Series Studies Contain Uncontrolled or Residual Confounding by Risk Factors for Acute Health Events?" *Regulatory Toxicology and Pharmacology* 51:135-140.

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(2):209-216.

Enstrom J. 2005. "Fine Particulate Air Pollution And Total Mortality among Elderly Californians, 1973-2002," *Inhalation Toxicology* 17:803-816.

Fewell Z, Davey Smith G, Stern JA. 2007. "The Impact of Residual and Unmeasured Confounding in Epidemiologic Studies: A Simulation Study," *Am. J. Epidemiol.* 166(6): 646-655.

Gamble J, Nicolich M. 2006. "Comments on the updated Harvard Six Cites Study," Correspondence, *American J. Respiratory and Critical Care Medicine* 174:722-724.

Goodman J. 2009. "Comments on the Epidemiology Evaluation in the Integrated Science Assessment for Particulate Matter Second External Review Draft," Gradient Corporation, on behalf of American Petroleum Institute (October 13).

Jerrett M, Burnett RT, Ma R, Pope CA, 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(6):1-10.

Krewski D, Burnett RT, Goldberg M, Hoover K, Siemiatycki J, Jerrett M, Abrahamowicz M, White WH. 2000. *Reanalysis of the Harvard Six Cities Study and the American Cancer Society Study of Particulate Air Pollution and Mortality.* Special Report, Health Effects Institute, Cambridge, Massachusetts (July).

Krewski D, Jerrett M, Burnett RT, Ma R, Hughes E, Shi L, Turner MC, Pope CA, 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.* Research Report #140, Health Effects Institute, Boston, Massachusetts (May).

Laden F, Schwartz J, Speizer FE, Dockery DW. 2006. "Reduction in Fine Particulate Air Pollution and Mortality, Extended Follow-up of the Harvard Six Cities Study," *Am J Respir Crit Care Med.* 173:667-672.

Lipfert FW, Baty JD, Miller JP, Wyzga RE. 2006. "$PM_{2.5}$ Constituents and Related Air Quality Variables as Predictors of Survival in a Cohort of U.S. Military Veterans," *Inhalation Toxicology* 18:645-657.

Lipfert F, Perry HM Jr., Miller JP, Baty JD, Wyzga RE, Carmody SE. 2000. "The Washington University-EPRI Veterans' Cohort Mortality Study: Preliminary Results," *Inhalation Toxicology* 12(Supplement 4):41-73.

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 U.S. Veterans," *Atmospheric Environment* 40:154-169.

McConnell 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," *Journal of Exposure Analysis and Environmental Epidemiology* 10:427-436.

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):11332-1141.

Smith AE, Chan NY. 1997. "How Statistics Can Mislead PM Policy: A Case of Smoke and Mirrors?" Attachment 4 to Comments of the Utility Air Regulatory Group on the National Ambient Air Quality Standards for Particulate Matter: Proposed Decision, 61 Fed. Reg. 65,638 (December 13, 1996) (Docket A-95-54) and Proposed Requirements for Designation of Reference and Equivalent Methods for $PM_{2.5}$ and Ambient Air Quality Surveillance for Particulate Matter: Proposed Rule, 61 Fed. Reg. 65,780 (December 13).

Villeneuve PJ, Goldberg MS, Krewski D, Burnett RT, Chen Y. 2002. "Fine Particulate Air Pollution and All-Cause Mortality within the Harvard Six-Cities Study: Variations in Risk by Period of Exposure," *Ann. Epidemiol.* 12(8):568-576.