Reassessing the Human Health Benefits from Cleaner Air

Tony Cox

tcoxdenver@aol.com

August 22, 2011

ABSTRACT

Recent proposals to further reduce permitted levels of air pollution emissions are supported by high projected values of resulting public health benefits. For example, EPA recently estimated that the 1990 Clean Air Act Amendment (CAAA) will produce human health benefits in 2020, from reduced mortality rates, valued at nearly two trillion dollars per year, compared to compliance costs of $65 billion (0.065 trillion) dollars. However, while compliance costs can be measured, health benefits are unproved; they depend on a series of uncertain assumptions. Among these are that additional life expectancy gained by a beneficiary (with median age of about 80 years) should be valued at about $80,000 per month; that there is a 100% probability that a positive, linear, no-threshold, causal relation exists between $PM_{2.5}$ concentration and mortality risk; and that progress in medicine and disease prevention will not greatly diminish this relationship. We present an alternative uncertainty analysis that assigns a positive probability of error to each assumption. This discrete uncertainty analysis suggests (with probability > 90% under plausible alternative assumptions) that the costs of CAAA exceed its benefits. Thus, instead of suggesting to policy makers that CAAA benefits are almost certainly far larger than its costs, we believe that accuracy requires acknowledging that the costs purchase a relatively uncertain, possibly much smaller, benefit. The difference between these contrasting conclusions is driven by different approaches to uncertainty analysis, i.e., excluding or including discrete uncertainties about the main assumptions required for non-zero health benefits to exist at all.

KEY WORDS: Clean Air Act Amendment 1990, air pollution health effects, uncertainty analysis, risk-cost-benefit analysis, Weibull distribution

Introduction

Media, regulatory, and advocacy reports and recommendations frequently suggest that particulate matter in outdoor air ($PM_{2.5}$) kills people and causes serious health problems. For example, an Environmental Protection Agency (EPA) web site warns that "Numerous scientific studies have linked particle pollution exposure to a variety of problems, including irregular heartbeat; nonfatal heart attacks; and premature death in people with heart or lung disease" (www.epa.gov/pm/health.html). Conversely, regulatory proposals to further decrease currently permitted levels of pollutants are increasingly supported by reference to large estimated or predicted health benefits from doing so. For example, in early 2011, the EPA released the results of its cost-benefit analysis of the 1990 Clean Air Act Amendments (CAAA). The assessment made two striking claims[1]: (1) As of 2020, the CAAA would produce estimated health benefits valued at approximately two trillion (i.e., two thousand billion) dollars per year, compared to estimated compliance costs of only about $65 billion per year; and (2) The uncertainties in the cost-benefit analysis are small enough so that, "***The extent to which estimated benefits exceed estimated costs*** and an in-depth analysis of uncertainties ***indicate that it is extremely unlikely the costs of 1990 Clean Air Act Amendment programs would exceed their benefits under any reasonable combination of alternative assumptions or methods identified during this study***" (emphasis in original).

This paper reexamines the assumptions, methods, and numbers used in reaching these encouraging conclusions. We come to the very different conclusion that EPA's evaluation of health benefits is unrealistically high, by a factor that could well exceed 1000, and that it is therefore very likely that the costs of the 1990 CAAA exceed its benefits, plausibly by more than 50-fold. Our reasoning involves re-examining specific uncertainties (including model uncertainty, toxicological uncertainty, confounder uncertainty, and uncertainty about what actually affects the timing of death in people) that were acknowledged qualitatively, but whose discrete contributions to uncertainty in health benefits were not quantified, in EPA's cost-benefit analysis.

Of greater methodological interest than the difference in conclusions is the difference in uncertainty analyses. If EPA's assessment of uncertainties were correct, then it should be impossible, or at least very improbable, for any plausible variations in assumptions to reverse the main conclusion that CAAA benefits greatly exceed its costs. Yet, this is precisely what we will show: plausible variations in assumptions easily reverse the sign of the estimated net benefits. The principal methodological challenge that we seek to address in this paper is, therefore, to diagnose the features of the uncertainty analysis that led to such a confident – but, we think, erroneous – conclusion, and to understand how to avoid such overconfidence in future risk-cost-benefit analyses.

1. The Main Logic of EPA's Benefit Assessment

Most of the benefit that EPA projects for 2020 is based on the estimated number and dollar value of mortalities that it expects will be prevented in that year by having cleaner air. The main calculation is simple:

Benefit = reduced number of deaths in 2020 x value per statistical life saved (1)

≈ (230,000 fewer deaths) x ($8,900,000 per life saved) ≈ $2 trillion.

The rationales for these two input values are provided in detail in EPA's report[2]. The 230,000 number for fewer deaths is given in Table 5-6, and is the estimated reduction in mortalities for all adults over 30 from reduced fine particulate matter ($PM_{2.5}$) in air. The value per statistical life saved (VSL) in 2020, of $8,900,000, is given in Table 5-4.The actual calculation is more refined, as it includes discounting and additional sources of benefits (e.g., from reduced morbidities and improved visibility, as well as reduced mortalities from lower ozone levels), but these make little difference to the final answer, which is still close to two trillion dollars worth of benefits for 2020, and even more in later years. As EPA states, "These avoided deaths are valued at 1.8 trillion ($2006), with primary low and primary high bounds on this estimate of $170 billion to $5.5 trillion."[2]. The reduction to $1.8 trillion reflects the passage of time before the estimated benefit is received. We will therefore focus on this simple calculation, and on the uncertainty analysis that led to these stated lower and upper bounds, as these calculations drive both the main quantitative results and the important qualitative conclusion that the value of health benefits almost certainly greatly exceeds the costs of compliance, which EPA estimates (Table 7-4) as $65 billion for 2020.

2. EPA's Input Estimates and Weibull Uncertainty Analysis

The upper and lower bounds for the two trillion dollar main estimate of benefits reflect uncertainties about the number of mortalities that will not occur in 2020 because of cleaner air, as well as uncertainty about the correct VSL value to use. The reduction in mortalities is estimated by multiplying estimated reductions in $PM_{2.5}$ concentrations by an estimated concentration-response potency factor, called the C-R coefficient. As EPA explains (p. 5-10), "We based the primary C-R coefficient estimate of the Second Prospective study on a Weibull distribution with a mean of 1.06 percent decrease in annual all-cause mortality per one $\mu g/m^3$. This mean is roughly equidistant between the results of the two most well-studied PM cohorts, the ACS cohort (0.58, as derived from Pope et al., 2002) and the Six Cities cohort (1.5, as derived from Laden et al., 2006), both of whose results have been robust to continued follow-up and extensive re-analysis." Likewise, for the VSL number, EPA again used a Weibull uncertainty distribution, noting (Table 5-4) that, "[The] Mean Value of Statistical Life (VSL) [is] based on 26 wage-risk and contingent valuation studies. A Weibull distribution, with a mean of $7.4 million (in 2006$), provided the best fit to the 26 estimates." The $7.4 million value is shown for 1990 income levels, and is updated to $8.9 million for 2020 income levels. These two Weibull distributions, for the C-R coefficient and VSL, together with estimated reductions in $PM_{2.5}$ concentration levels due to CAAA, drive the 2 trillion dollar main benefit estimate and the upper and lower bounds around it.

3. Questioning the Logic: Does the Main Calculation Make Sense?

EPA's main analysis, as just summarized, has the commendable virtues of simple, transparent logic and well documented input values and reasoning. It invites open inspection by the scientific community, so that any important deficiencies can be noted and, if necessary, repaired. To such deficiencies, we now turn.

The logic of equation (1) requires estimating the reduction in mortalities in 2020 due to cleaner air (reduced $PM_{2.5}$ concentrations) resulting from the CAAA. The total number of mortalities in 2020 *with* the CAAA is (or should be) calculated as follows: From the number of mortalities that would have occurred in 2020 without the CAAA (as estimated from population sizes and mortality rates), *subtract* the number of mortalities that are postponed until 2021 or later because of reduced mortality rates attributed to the CAAA; and *add* the number of mortalities that are postponed until 2020 from 2019 or earlier because of reduced mortality rates attributed to the CAAA. The difference between this number (i.e., the mortalities that occur in 2020 *with* the CAAA) and the original number (i.e., the mortalities that occur in 2020 *without* the CAAA) is the total change in mortalities in 2020 attributable to the CAAA. In other words,

deaths in 2020 <u>with</u> CAAA = number that would have occurred <u>without</u> CAAA – deaths postponed because of CAAA until after 2020 + deaths postponed because of CAAA until 2020 from earlier (2a)

Similarly,

Change in deaths in 2020 attributable to CAAA = – deaths postponed because of CAAA until after 2020 + deaths postponed because of CAAA until 2020 from before 2020 (2b)

Unfortunately, EPA's calculation includes only the first term in the change (that is, the reduction – *deaths postponed because of CAAA until after 2020*), but neglects the second term (that is, the increase *deaths postponed because of CAAA until 2020 from before 2020*). EPA estimates that the average gain in life expectancy is less than a year (Table 5-8), so if mortality rates and population sizes change relatively slowly between 2019 and 2021, the deaths postponed from 2020 to 2021 or later will be approximately equal to the deaths postponed until 2020 from 2019 or earlier. Hence, *the net change in deaths in 2020 will be approximately zero,* instead of the 230,000 reduction in mortality estimated by EPA.

More generally, calculating "deaths avoided" in any year is not the correct risk analysis tool to use to estimate the health benefits from a regulation or other intervention. A correct analysis (e.g., using age-specific hazard functions and life table analysis) shows that even a regulation that successfully decreases age-specific mortality rates, thereby increasing life expectancy, may not reduce the annual number of deaths in the population.

Example: Accounting for Deaths

The following simplified numerical example illustrates why a regulation that extends lives does not necessarily reduce deaths per year. For simplicity, suppose that, before the regulation, each individual lived to age 80, and then had a 50-50 chance of either dying at age 80, or of surviving for a further decade and dying at age 90. Suppose that the regulation prevents early deaths, so that all those who would have died at age 80 now die at age 90. (Those who would have died at age 90 still do so.) Now,

how many people die each year before and after the regulation? (Assume a constant birth rate of *N* people per year, which is unaffected by the regulation.) Before the regulation, *N* people die each year (half of the cohort of *N* people born 80 years ago + half of those born 90 years ago, for a total of *N*/2 + *N*/2 = *N*.) After the regulation has fully taken effect, *N* people still die each year (all of the cohort born 90 years ago). Although half of the population lives longer, the number of deaths per year is unchanged.

Of course, even if the net deaths prevented in 2020 are zero (as will necessarily be true under steady-state conditions, and as will be approximately true if the population size is changing relatively slowly on a time scale of months to years), this would not imply that increasing life expectancy by several months has no value. But it does suggest that equation (1)'s emphasis on reduced mortalities in 2020 is the wrong way to capture this value. The problem is that the main projected health benefits are not (or should not be) deaths "prevented" or "avoided," or lives "saved," but rather life-years gained. EPA's own report[2] explains this well: "Avoided premature mortality is one of the more commonly cited results of benefits analyses for air pollution control. However, as noted in the valuation section of this chapter, *a more accurate description of the benefit of clean air is a reduction in the risk of mortality* for the exposed population over many years, which results in the extension of lives (sometimes referred to as 'lives saved'). *Other useful metrics of the benefit of cleaner air are the number of life years that are gained through the reduction of mortal risks*, and the number of years of life expectancy gained on average throughout the population." (Emphases added.)

Accordingly, we can re-do the calculation of benefits in equation (1), switching from deaths "prevented" or avoided (which, in reality, do not occur, since reducing $PM_{2.5}$ does not confer immortality) to life-years gained. EPA estimates (Table 5-8 of [2]) that, in 2020, almost two million (1,900,000) life-years will be gained because of cleaner air

caused by CAAA. No value of a statistical life-year ($VSLY) gained is explicitly identified in EPA's report, and values in the literature vary widely, depending on the elicitation techniques used and the ages of the recipients. We tentatively use a $VSLY value of $150,000 per life-year, which has previously been used for life-years gained by cancer patients through the year 2020[3]. (For comparison, if we assumed, simplistically, that the $8,900,000 VSL number used in equation (1) consists of at least 50 equally-valued life-years, then the corresponding $VSLY value would be at most $8,900,000/50 = $178,000. This ignores the important complexities of age-specific $VSLY values and discounting, but suggests that the estimated $VSLY value of $150,000/year is not grossly inconsistent with EPA's $VSL number.) Equation (1) can now be replaced with the following estimated value of health benefits based on estimated life-years gained:

Benefit = (life-years gained in 2020) x ($VSLY per life-year gained) (3)

= (1,900,000 life-years gained) x ($150,000/life-year gained) = 0.3 trillion.

Thus, simply replacing lives saved with life-years gained, and hence the $VSL value in equation (1) with the $VSLY value in equation (3), reduces the benefit estimate from about 2 trillion to about 0.3 trillion.

As it happens, a similar numerical conclusion results even if the logic of equation (1) is retained, but the ages at which projected benefits occur are taken into account. EPA estimates the median age of people whose lives are "saved" (i.e., who gain extra months of life) from cleaner air to be close to 80 years (Table 5-8 of [2]). The $VSL of $8,900,000 is appropriate for a healthy young adult of age about 25 (Aldy and Viscusi[4], Figure 1). The $VSL for people aged close to 80 has been estimated to be

approximately one sixth of the $VSL for people aged about 25 (Murphy and Topel[5], Figure 1). Thus, using $VSL numbers that condition on the ages at which benefits are received would also reduce the estimated value of health benefits substantially, with a six-fold reduction reducing EPA's current estimate of $1.8 trillion to a revised estimate of 0.3 trillion. Despite this numerical coincidence, however, we believe that equation (3) is logically (and perhaps ethically) preferable to equation (1), since, as EPA notes, the true benefit from reduced mortality rates is gain in expected life-years.

Two other ways of framing the simple analysis in this section may help to clarify the crucial role that conditioning on age should play in health effects benefits assessments. The first is that, in very round numbers, about two million life-years gained in 2020 can create about two trillion dollars worth of benefit only if each life-year gained is valued at about a million dollars (or, equivalently, at about $80,000 per additional month of life). This is a much higher value than many octogenarians place on their own gains in life-years[4,5]. Second, applying a $VSL value appropriate for a 25-year old to people who are mostly over 75 (and a significant fraction of whom are over 100 years old, according to EPA's[2] Table 5-8, since this group contributes as much to estimated life-years gained as the 40 to 49 year olds in 2020) suggests that cleaner air is being evaluated, perhaps inadvertently, as a sort of fountain of youth, restoring the full value of youth (e.g., with at least another 50 years of healthy life expectancy) to people who, in fact, have a median age close to 80 years (*ibid*). This appears to be overly optimistic.

Health economics studies of the costs and benefits of medical interventions routinely quantify health benefits in terms of the monetized value of quality adjusted life

years (QALYs), per the recommendation of the U.S. Panel on Cost Effectiveness[19]. QALY values of between $50k and $100k per year are common, and values up to $200k can be found, but these values are substantially less than the $1 million that EPA implicitly places on a life-year (without adjusting for quality). Thus, whether one uses monetized values of "lives saved," or life-years added, or quality-adjusted life-years added, EPA's valuation appears to be higher than others, for similarly aged (and perhaps similarly infirm) people, by approximately one order of magnitude.

4. Questioning the Major Premise: Will Reducing $PM_{2.5}$ Concentrations Further Really Reduce Mortality Rates Further?

Even more important than uncertainty about the correct value of $VSL (or of $VSLY) is uncertainty about the correct value of the C-R coefficient that converts reductions in $PM_{2.5}$ concentrations to corresponding projected reductions in mortality rates. If the C-R coefficient is zero, then the health benefit from further reducing $PM_{2.5}$ concentrations is also zero. The assumption that the C-R coefficient is positive might be wrong. EPA's qualitative discussion of key uncertainties (Table 5-11 of [2]) explicitly identifies this crucial assumption, noting that the "Analysis assumes a causal relationship between PM exposure and premature mortality based on strong epidemiological evidence of a PM/mortality association. However, epidemiological evidence alone cannot establish this causal link." The possible impact of this assumption is identified as being a "potentially major" overestimate of benefits. However, EPA rates its confidence in its causal hypothesis as "High," arguing that, "The assumption of causality is suggested by the epidemiologic and toxicological evidence

and is consistent with current practice in the development of a best estimate of air pollution-related health benefits. At this time, we can identify no basis to support a conclusion that such an assumption results in a known or suspected overestimation bias." EPA's quantitative uncertainty analysis goes further, implicitly expressing 100% confidence that this causal hypothesis is correct. The Weibull uncertainty distribution for the C-R coefficient implicitly assigns a probability of 100% to the C-R coefficient being positive, and hence a 0% probability to it being zero (or negative). Thus, the qualitative assessment of "high" confidence is implicitly mapped to a quantitative assessment of 100% certainty. Such complete certainty is unwarranted by available data and knowledge, as discussed next.

Is There Really a Significant Positive Association Between $PM_{2.5}$ and Mortality Rates?

Being unable to identify a basis for a suspicion of an overestimation bias does not, of course, guarantee that there is no such bias. Nor, perhaps, should a (subjectively perceived) suggestion of the possibility of a causal relation in epidemiological and toxicological data, or consistency with current practices in developing best estimates, warrant a very confident conclusion – let alone 100% confidence – that a causal relation necessarily exists.

For toxicological data, both theory and empirical evidence indicate that the inflammation-mediated adverse health effects of fine particulates probably have thresholds, below which the C-R coefficient is zero[6,7]. For example, low levels of exposure induce increased antioxidant (AOX) production in the lung, but higher levels induce levels of reactive oxygen species (ROS) that overwhelm the very limited capacity

of the lung to increase AOX. Increased disease risks occur only when there is an imbalance that leads to a sustained high-ROS lung environment and other changes that disrupt homeostasis, which may not occur at or near currently permitted exposure levels[6].

For epidemiological data, the National Academy of Sciences and several expert statisticians have long questioned whether reported findings of significant positive C-R coefficients may be misinterpreting ambiguous data as showing evidence of a significant association where none actually exists[8-10]. This concern is based largely on awareness that multiple testing biases (which inflate the rate of false positives) can arise when pollution health effects investigators make numerous modeling choices (e.g., of which pollutant summaries, confounders, covariates, and interactions to include in the model; what logarithmic or other transformations to apply to them; lags in variables; model form specifications; treatment of exposure uncertainties, etc.) in arriving at the reported significant positive associations. Attempts to increase the objectivity of conclusions about PM-mortality rate associations, by considering all of the effect estimates from many (e.g., thousands of) computer-generated models that fit the available data approximately well, without applying human judgment to select or defend any particular model or conclusion, have found that the hypothesized associations may not exist. For example, some Bayesian model-averaging (BMA) analyses (averaging effects over multiple models, weighted by their likelihoods in light of the data) of time series studies have reported, contrary to findings based on subjective selection of modeling assumptions, that available time series data sets give no overall indication of a positive relation between $PM_{2.5}$ and mortality rates[8, 9]. The probability that current or

lagged $PM_{2.5}$ values help to predict mortality rates (compared to predictions based solely on non-pollutant variables such as temperature and pressure) ranges from about 2% to 7% in some BMA studies[9], and associations between pollutants (including particulate matter and ozone) and mortality rates have been reported as not significantly different from zero in some Canadian and U.S. investigations[8-10]. However, the BMA analyses were conducted for time-series data, and thus are not directly comparable to the long-term cohort studies used by EPA. We do not know whether the conclusion that only a small fraction of models show any association of PM with mortality rates would hold in the two main studies (the Harvard Six Cities and American Cancer Society cohort study data) used in EPA's assessment[2].

Against such reports of apparently non-existent associations, some investigators whose methods and models have informed EPA's benefits assessment have argued that "Bayesian approaches in general and BMA in particular are not panaceas for model selection." (This is certainly true, as they are not model selection methods at all.) They further note that Bayesian approaches may be mistakenly applied, and therefore urge that "model selection should be informed by substantive knowledge about the environmental health processes influencing the outcome."[11] This response appears to reject the BMA principle that model selection should not be used at all (since even the "best" single model is almost certain to be wrong); but to advocate instead the continued use of judgment, "informed by substantive knowledge," as a basis for selecting models (which, in turn, may drive conclusions that are more consonant with the modelers' judgments about what should be found) – thus reopening the possibility that confirmation bias (selecting modeling assumptions that are favored by prior

expectations or expert opinions) and multiple selection bias might explain some reported positive associations. For example, imposing an *a priori* modeling assumption that mortality rates increase approximately in direct proportion to pollutant levels at low concentrations (e.g., in a Cox proportional hazards model), or that the C-R function is linear at low concentrations (and piecewise linear overall), makes it possible to conclude that the C-R curve "is linear, clearly continuing below the current U.S. standard of 15 $\mu g/m^{3}$"[12]. In this way, substantive knowledge or *a priori* assumptions and model selection constraints can still be used to produce low-dose linear results and a positive C-R coefficient. EPA (2011, Table 5-11, p. 5-42) identifies the "Assumption of a linear, no-threshold model for PM and ozone mortality" as a "Key uncertainty," but assesses as "High" its subjective confidence that this choice of model is correct, and as "Probably minor" the potential impact on overestimation of benefits of considering other model forms. This contrasts with time series studies that conclude that considering other models removes all significant positive associations between $PM_{2.5}$ and mortality rates[8-10]. A possible explanation is that the one cohort study that EPA's table cites in support of these crucial conclusions (an EPA- (Health Effects Institute)-funded study of the ACS data[24]) focused on models (random effects Cox proportional hazards and Poisson regression models) that imply low-dose linearity, rather than considering other (e.g., J-shaped or threshold) models that could have revealed absence of a positive C-R relationship at low exposure levels.

Even without further discussing, or seeking to resolve, how best to use computationally intensive statistical methods such as BMA to more fully characterize uncertainty about the C-R coefficient, it seems clear that there is significant uncertainty

about whether a true association exists between ambient $PM_{2.5}$ concentrations and mortality rates; it is still being investigated. This crucial uncertainty should not be lightly dismissed. EPA's qualitative caveat deserves emphasis, that the "Analysis assumes a causal relationship between PM exposure and premature mortality based on strong epidemiological evidence of a PM/mortality association, [but] epidemiological evidence alone cannot establish this causal link." Even this caveat should perhaps be tempered further, by acknowledging that calling the epidemiological evidence favoring a PM/mortality rate association "strong" may largely reflect unvalidated modeling assumptions (e.g., of a linear, no-threshold C-R relation at ambient levels). Frequent minority reports of statistically significant *negative* C-R coefficients, in models that allow this possibility, raise further questions about the coherence and reality of claimed statistical associations. For example, a recent review of associations between $PM_{2.5}$ and all-cause and cause-specific mortality in 27 U.S. communities indicated *reduced* all-cause mortality rates at increased levels of $PM_{2.5}$ for one third of the communities, including Birmingham, Dallas, Houston, Las Vegas, and Riverside[11].

To understand the sensitivities of EPA's cost-benefit assessment to plausible alternative assumptions, it suffices to note that the question of whether a real (data-driven, rather than model selection-driven) association exists between $PM_{2.5}$ concentrations and mortality rates remains open, apparently on fairly substantive grounds, with at least some BMA studies indicating rather strongly (e.g., with greater than 95% probability) that there is no such association[9]. Lack of consistency in effects across studies, and large, unexplained heterogeneity in apparent C-R functions (including some with negative slopes) across locations, add qualitative reasons to

question whether a stable, coherent association exists[8-10,13]. This uncertainty should be explicitly addressed and, if possible, quantified, as BMA analyses seek to do, rather than being tacitly ignored, e.g., by using only a single average value across all studies[13] (perhaps with subjective confidence intervals) without explaining or resolving the observed negative relations.

Are Reported Statistical Associations Causal?

Even if a real (model-independent) statistical association exists, of course, it might not be causal. Many investigators have recognized that covariates and confounders (such as proportion of high-risk minorities in a city, or correlates of wealth or education or poverty, or residual confounding by weather variables, including high temperatures or barometric pressures, or changes and trends in these variables) might create significant statistical associations between pollution and mortality rates, even if the former does not necessarily cause the latter. Likewise, regressing some trend variables (e.g., declining cause-specific and overall mortality rates in the wake of innovations such as angioplasty, anti-smoking campaigns, and aspirin therapy) against others (such as declining ambient pollution levels in recent decades) can show strong, statistically significant associations between them, with reductions in exposure concentrations routinely being followed (as well as preceded) by reductions in mortality rates, even if there is no causal relation between them.

Such threats to valid causal interpretation of statistical associations are well recognized by many investigators in pollution health effects research, but approaches for trying to address them have varied widely. To meet such challenges and threats to

valid causal inference more decisively, econometric and statistical tests for potential causation between variables have been extensively developed, using concepts such as conditional independence and Granger causality for multiple time series[14]. However, these methods have generally not yet been applied to $PM_{2.5}$ and mortality data. Thus, the question of whether any statistical associations might be causal has not been settled using such relatively formal, objective tests. (As previously noted, inconsistencies across cities and studies, and large, unexplained heterogeneity in C-R coefficients estimated in different studies, including some significant negative ones, undermine the coherence of proposed causal interpretations. Thus, the question of whether there is any positive *causal* relation between ambient levels of ozone and $PM_{2.5}$ and mortality rates, is a second important, discrete, uncertainty that should be explicitly addressed in uncertainty analysis. It cannot be answered by arguing that there is a positive *statistical* association between them (at least under some model-selection assumptions), since statistical and causal relations have no necessary connection (e.g., omitted confounders can make statistical relations positive even if causal relations are zero or negative). Nor can it be settled by concluding, based on expertise, that the data are *consistent* with such a relationship[2], if the data are at least equally consistent with the opposite conclusion of no relationship[8-10]. We conclude that the question remains unsettled, and that it is appropriate for inclusion in uncertainty analysis.

EPA's benefit assessment[2] (pp 5-7 to 5-10) recognized and addressed the possibility of confounding in its uncertainty analysis, as follows:

"Several other sources of uncertainty exist in the relationship between ambient pollution and the health outcomes, including model uncertainty, potential confounding by factors that are both correlated with the

health outcome and each other, and potential misclassification of the study population exposures. …EPA recently conducted an expert elicitation (EE) study, which is the formal elicitation of subjective judgments, in order to more fully characterize the uncertainty surrounding the $PM_{2.5}$/mortality C-R function. This study allowed experts to consider and integrate several sources of uncertainty in the form of a probability distribution for the C-R function…[W]e rely exclusively on longterm studies to quantify PM mortality effects. This is because cohort studies are able to discern changes in mortality rates due to long-term exposure to elevated air pollution concentrations. …Based on consultations with the Council's Health Effects Subcommittee (HES), the 812 Project Team developed a distribution of C-R function coefficients (i.e., the percent change in annual all-cause mortality per one $\mu g/m^3$ change in annual average $PM_{2.5}$) for use in the PM-mortality C-R function for the Second Prospective study. This distribution is rooted in the epidemiological studies that most inform our understanding of the PM mortality C-R function, but reflects the broader findings of the EE study. We based the primary C-R coefficient estimate of the Second Prospective study on a Weibull distribution with a mean of 1.06 percent decrease in annual all-cause mortality per one $\mu g/m^3$. This mean is roughly equidistant between the results of the two most well-studied PM cohorts, the ACS cohort (0.58, as derived from Pope et al., 2002) and the Six Cities cohort (1.5, as derived from Laden et al., 2006), both of whose results have been robust to continued follow-up and extensive re-analysis."

Thus, EPA explicitly recognized the possibility of confounding. They dealt with uncertainty about whether confounding, model uncertainty, or other sources of non-causal associations (such as exposure measurement error or misclassification) might have created an overall false-positive association between $PM_{2.5}$ concentrations (C) and health responses (R) in past studies, by (a) Relying on two studies that reported significant positive associations; (b) Discarding, or leaving unexplained, studies that reported significant negative or non-significant C-R associations; and (c) Applying subjective expert judgments, encoded as a Weibull distribution (which cannot show zero or negative relationships), to conclude that a positive causal relation exists with 100% subjective confidence probability. (The Weibull distribution implicitly assigns a subjective probability of zero to the discrete possibility that fully controlling for all

relevant confounders and trends would yield a negative or zero causal estimate of the percent decrease in annual all-cause mortality per μg/m^3 of $PM_{2.5}$.)

Whether or not their conclusion is correct, this *process* – considering only data and probability distributions that support or logically require, *a priori*, the conclusion of a positive relation between exposure and health effects – is surely not adequate for characterizing uncertainty about whether the underlying assumption that the true (unconfounded) C-R relation is positive is correct. The results of expert elicitation, expressed in the form of a distribution for a (causally interpreted) C-R coefficient, are also of uncertain value as a guide to objective truth, insofar as (a) The elicited distribution does not match, and hence does not explain, the empirically observed distribution of C-R coefficients in meta-analyses of dozens of reputable studies, about a third of which are negative[11]; and (b) No one currently knows how or whether further reducing ambient concentrations of $PM_{2.5}$ *per se* – as opposed to reducing other activities, conditions (e.g., temperature extremes), and pollutant constituents (e.g., various oxides and sulfates) that are strongly correlated with $PM_{2.5}$ – would cause any incremental reductions in mortality risks.

Experts, like other people, may have opinions on these matters. But one cannot elicit from experts knowledge that they do not have. In this case, correct "substantive knowledge" of causal relations showing how reducing current and future ambient $PM_{2.5}$ concentrations would affect human health, or how and whether other confounders and their interactions affect the estimated C-R relations for $PM_{2.5}$ is not yet available. We do not yet even know which components of $PM_{2.5}$ (if any), possibly interacting with which other pollutants, cause adverse health effects under present ambient exposure

conditions. This does not necessarily imply that no new action should be taken. But it does imply that, if $PM_{2.5}$ *per se* is not involved in causing increased mortality rates, then further reductions in ambient $PM_{2.5}$ levels may not produce any health benefits. This possibility should be acknowledged in uncertainty analysis of human health benefits by including a substantial discrete (non-zero) uncertainty about causation.

Do C-R Relations Have Thresholds or J-Shapes?

Even if historical data showed a clear, unambiguous, causal relation between concentration and response, with past declining levels of $PM_{2.5}$ (or C) consistently causing proportionally declining levels of mortality rates (or R), it would be important to establish whether further reductions in C would produce further reductions in R. At some point, C might reach (or might already have decreased below) a threshold below which no further reductions in R are gained. Uncertainty about that threshold, and about whether current standards are already below it, should then be quantified. Or, conceivably, continuing reductions in C might eventually pass a point after which further reductions in C actually increase R, as suggested by a few of the U-shaped or J-shaped empirical C-R model curves fit to PM_{10} data for the twenty largest U.S. cities[15]. (This might also explain the previously noted disconcerting finding of negative C-R coefficients in approximately a third of cities examined[11].) The probability of this contingency, and the conditional probability that we have already approached or passed the bottom of such a J shape if it exists (at least in some cities), should limit the certainty with which additional reductions in C can be projected to cause additional reductions in R.

Existing discussions in the air pollution health effects literature of the possibility of a threshold for the C-R function largely recapitulate the technical issues in discussions of whether a non-zero C-R coefficient exists at all. Those who favor a linear, no-threshold model argue that model selection criteria such as the Akaike Information Criterion (AIC) can be used to support it[12], insofar as other models tested typically do not give significantly better AIC scores. Those who are more agnostic about the linear no-threshold model point out that the AIC criterion is not appropriate for this use: in simulation experiments (for which the correct model is known) it only identified the correct model about half the time, and assuming a linear no-threshold model can greatly under- or over-estimate true mortality risks if non-linearities are present but are mistakenly assumed away based on AIC[16]. Defenders of the linear no-threshold model further counter that population data exhibit no thresholds, but appear to be linear down to the lowest levels measured[15]. Skeptics could respond that this pattern is precisely what one might expect if there really is a population threshold, but individual exposure estimates contain unmodeled errors, as in past studies; thus, the observed population C-R function is irrelevant to determining whether there is a true C-R threshold[6]. That errors in past exposure estimates cast doubt on inferences about health effects is starting to be more widely acknowledged[17], but no one has yet advanced a model that clearly fits most of the data, reconciles past inconsistencies, accounts for past measurement errors, and resolves the debate with an obviously correct C-R relation.

As in the case of uncertainty about whether a positive C-R coefficient exists, the role of uncertainty analysis is not to resolve whether threshold or J-shaped C-R relations exist (at at least some locations), but only to make sure that any significant

uncertainties about these possibilities are identified and explicitly addressed in the presentation and evaluation of uncertain health benefits projected from reductions in exposure concentrations. However, EPA's quantitative uncertainty analysis implicitly assumes that there are *no* uncertainties about the correct form (linear, no-threshold with probability 100%) or sign (positive with 100% probability, zero or negative with 0% probability) of the C-R relation. Admitting that there are such uncertainties would substantially change both the uncertainty analysis and the evaluation of projected health benefits, as explained in Section 5.

How Much Smaller Will Future C-R Coefficients be than Past Ones?

Finally, suppose that a true, causal, linear C-R coefficient were to be confidently identified in historical data. Even this would not warrant a conclusion that *future* reductions in pollutant levels would achieve proportionate reductions in mortality rates and gains in life-years. Progress in medicine, including preventive medicine, is steadily shrinking the pool of highest-risk individuals whose lives might, under the above assumptions, be most prolonged by reducing pollutant levels. Long-term trends such as improved diet and exercise among adults, greater use of low-dose aspirin regimens and other medications to successfully prevent cardiovascular disease events, smoking reduction and cessation programs, and new treatment and therapeutic options (such as angioplasty) have led to a long-term trend of declining mortality risks from key diseases identified by EPA as being reduced by lower pollutant levels.

Some experts interpret epidemiological data as showing that "Short-term and long-term studies clearly indicate that relatively modest exposures to particulate matter in the ambient air are associated with increased morbidity and mortality due to coronary heart disease."[20] Death from coronary heart disease (CHD), especially among the elderly and patients with pre-existing cardiovascular diseases, has been proposed as a specific mortality risk caused by high levels of particulate pollution[20, 21]. The future relation between $PM_{2.5}$ and mortality risks, therefore, may be affected by trends in CHD risk. These trends show very significant declines in CHD mortality risks over the past four decades, with further declines expected. For example, a recent review[22] that quantified temporal trends in cardiovascular mortality risks in Minneapolis-St Paul, MN found that mortality rates for non-hospitalized men fell by about 2/3 between 1985 and 2008 (from about 150/100,000 in 1985 to about 50/100,000 in 2008). Similarly, for the United States as a whole, a recent World Health Organization (WHO) study[23] found that "Since the late 1970s, age-adjusted CHD mortality rates have been halved in most industrialized countries, including the United States. …Approximately 44% of the substantial CHD mortality decline in the United States between 1980 and 2000 was attributable to changes in major risk factors [smoking, systolic blood pressure, total blood cholesterol, body mass index], and 47% to specific cardiological treatments. These findings resembled those from other industrialized countries. …In conclusion, implementing evidence-based policies to better control tobacco use and achieve a healthier diet across the population could potentially halve future CHD deaths in the United States." Although future reductions in risk will depend on future trends in diet, exercise, obesity, aging of the population, and medical interventions and treatments,

there appear to be continued large opportunities for continued risk reductions in the United States (*ibid*). As such trends continue, the incremental benefits from additional pollution control might also be expected to shrink, as fewer people remain at high risk of mortality from pollution-associated diseases such as CHD. Thus, analysis of projected future benefits should include uncertainty about the size of future C-R coefficients, rather than assuming that C-R coefficients estimated from historical data will remain undiminished by current and future trends.

5. Discrete vs. Continuous Uncertainty Analysis

The major uncertainties about the health effects of reducing pollutant concentrations in air can be summarized as the following series of questions.

- *Is there a true C-R association?* What is the probability that there is a true (independent of model selection assumptions) statistical association between lower pollutant concentrations (C) and lower mortality rates (R)?
- *Is it causal?* Assuming that there is a true C-R association, is it causal? That is, will future reductions in $PM_{2.5}$ and ozone cause corresponding reductions in future mortality rates? Or, to the contrary, is the association non-causal, e.g., reflecting past downward trends in both variables that do not represent a causal relation, and that do not support manipulating future mortality rates by changing future ambient pollutant concentrations? What is the probability that the C-R association (if there is one) is causal?

- *Is it linear non-threshold (or are present and future concentrations above any applicable threshold, or on the ascending part of any J-shaped C-R function)?* Assuming that a causal C-R relation exists in past data, for concentration levels prevalent then, how likely is it that a similar C-R coefficient will continue to hold in future, as ambient concentrations continue to fall? What is the probability that we are *not* yet at or below a nadir, threshold, or threshold-like nonlinearity, in the C-R function, which would cause future reductions in C to produce no further reductions in R, and hence no incremental health benefits? Colloquially, we might wonder: How clean is clean enough to achieve the possible health benefits, and have we already reached a point past which no additional health benefits will accrue to additional reductions in concentrations? What is the probability that this is not the case?
- *Is it stable over time?* What is the probability that C-R coefficients will remain approximately the same, despite improvements in prevention and medical treatments for the diseases that are associated with pollutants? If the C-R coefficient does decline in future, as the at-risk population declines, then what is its new expected value?

A notable aspect of these uncertainties is that they are discrete: one can imagine answering yes or no to each one, if perfect information were available, and can therefore envision assigning a discrete probability (a number between 0 and 1) to each of these possibilities now, based on the imperfect information available now.

For purposes of plausible numerical illustration and sensitivity analysis, suppose that the quantitative probabilities for answers to these questions are as follows:

- *Probability that there is a true association* = 50%. As discussed previously, the data on associations are conflicting, with both positive and negative significant C-R associations being reported, and with results of individual studies appearing to be sensitive to choices of statistical methods and modeling assumptions[8]. Bayesian Model Averaging (BMA) analyses that permit the possibility of no association between $PM_{2.5}$ and mortality rates generally find little support for the hypothesis that such an association exists at ambient concentrations; the probability that it does exist was estimated as between about 2% and 7% in a recent BMA analysis[9]. Against this conclusion, many experts are convinced that the combined findings from multiple epidemiological studies "clearly indicate that relatively modest exposures to particulate matter in the ambient air are associated with increased morbidity and mortality due to coronary heart disease,"[20] and that the main remaining scientific challenge is to figure out why and how this occurs. It is not necessary to choose between these conclusions for purposes of illustrating how such uncertainties affect final benefits estimates. Instead, we will use a 50% probability of a true association as a point of departure for quantitative analysis, and then use sensitivity analysis to understand how changes in this probability affect conclusions.
- *Probability that a true association, if it exists, is causal* (and not explained away by trends, confounding, or the distinction between statistical and causal associations) ≤ 0.5. The inconsistencies and large unexplained differences in

estimated C-R functions, signs, and magnitudes across studies – even for the same chemical, such as ozone – and failure to find clear evidence of a coherent causal relation at ambient levels in clinical, laboratory, or experimental studies despite decades of evidence might suggest a "more likely than not" conclusion that there is not a causal relation. A probability interval of ≤ 0.5 captures this uncertain probability. This is strengthened by some past reviews and conclusions, e.g. that, "Because the mortality risk estimates from important observational epidemiologic studies are extremely weak, derived from studies unable to control for relevant confounding causes, and inconsistent by location, toxicologic and clinical information is necessary to judge the likelihood and degree to which such findings are causal. Toxicologic data on typical forms of pollution-derived PM strongly suggest that current ambient concentrations in the U.S. are too small to cause significant disease or death. …The expectation that lives will be saved by reducing ambient $PM_{2.5}$ in the U.S. is not supported by the weight of scientific evidence, although other bases for regulating PM may be justifiable."[18]

- *Probability that currently permitted concentrations are above any threshold* or nadir in the C-R function (if any; or probability that there is no such threshold or nadir) ≤ 0.5. Repeated finding of negative as well as positive C-R coefficients might suggest the possibility of a J-shaped function, and other reasons for expecting a threshold have previously been discussed[6]. The inconsistencies and difficulties in identifying any clear positive C-R association[18] also suggest that we could now be below the clearly rising portion of a C-R function.

- *Expected reduction factor in C-R coefficient by 2020* ≤ 0.5, assuming that a linear no-threshold causal C-R coefficient exists. Improved medical diagnosis, intervention (e.g., with aspirin therapy, smoking cessation programs, or cholesterol-reducing measures, including changes in diet and exercise), and treatment, which are increasingly being used, can potentially reduce premature CVD mortalities by 80%-90% or more, but an aging or more obese population over the next decade could modestly increase the C-R slope[23]. Based on these opposed trends, and in light of the past four decades of declining CVD risks, an overall reduction factor of 0.5 might not be implausible as a starting point for the reduction in any real C-R coefficient, although higher or lower values might be developed with additional modeling. We will use 0.5 as a point of departure for understanding the sensitivity of estimated benefits to this factor.

Each of these factors is conditioned on the outcomes of all of its predecessors being such that a positive linear no-threshold causal C-R coefficient exists.

With these rough estimates of factors (using the upper end of each uncertainty interval, e.g., 0.5 for the interval ≤ 0.5), the health benefits estimated by EPA would be adjusted as follows:

(1.8 trillion initial estimate) x (1/6 reduction factor for $VSL if age or $VSLY is considered) x (0.5 probability that a true association exists) x (0.5 probability that a true association is causal, given that one exists) x (0.5 probability that ambient concentrations are above any thresholds or nadirs in the C-R function, given that a true causal C-R relation exists) x (0.5 expected reduction factor in C-R coefficient by

2020 due to improved medication and prevention of disease-related mortalities) = (1/6)*0.5*0.5*0.5*0.5 = $19 billion.

In round numbers, the combination of assumptions shown here reduces the estimated value of health benefits in 2020 one hundred-fold, from two trillion dollars to $20 billion. Instead of being much greater than the estimated compliance costs of $65 billion, the estimated value of health benefits is substantially less. This conclusion is robust to changes in any of the above factors, in that replacing any of them with 1 (no reduction) would not change the conclusion that benefits are less than costs. Sensitivity analysis of this product is trivial: changing any factor by *k*-fold changes the entire product by *k*-fold, for any $k > 0$.

The point of departure values for this calculation were deliberately chosen realizing that they might be conservative (erring in the direction of over-stating C-R relations and health benefits), with all factors ≤ 0.5 being rounded up to 0.5 for purposes of this initial calculation. Two main changes that might increase realism would significantly reduce estimated benefits. One would be to accept at face value the BMA results indicating that a true positive association between present ambient levels of $PM_{2.5}$ and mortality rates is unlikely[9]. This could reduce the point of departure for this probability by about a further order of magnitude, from 0.5 to about 0.05, and hence reduce the plausible high estimate for monetized value of annual health benefits from $19 billion to $1.9 billion. (A strong conviction that causality is very unlikely would have a similar effect.) The other main change would be to replace the 1/6 reduction factor for $VSL with a reduction factor of 1/10 for $VSLY (or value of a QALY), as discussed

earlier. This would reduce annual benefits estimates by a further factor of 0.6 (e.g., from $1.9B to approximately $1B). Such numbers are well below the $65B in annual compliance costs estimated by EPA. If these two alternative values were used as points of departure, then the conclusion that expected benefits are less than expected costs would be robust to changes that increased any or all of the remaining factors to 100%.

These calculations are all based on expected values (albeit for upper bounds, if the points of departure are indeed conservative). For a risk-averse individual or society, however, the economic value of an uncertain benefit is *less* than its expected value, by an amount that depends on the risk premium for uncertainty. A further reduction in the evaluation of the uncertain benefits (by perhaps a further factor of 2, depending on the risk premium) would be needed to take into account risk aversion for uncertain gains.

Such further adjustments and refinements, while possibly producing a more realistic answer, might also invite inessential controversy. Our main conclusion, that the 1990 CAAA benefits plausibly amount to only a small fraction of their direct compliance costs (the only costs considered in the $65 billion estimate), does not depend on choosing very small probabilities as points of departure, nor on greatly revising EPA's value judgments about $VSL or $VSLY, nor on incorporating risk aversion for uncertain benefits. Rather, it is a robust result of the combination, in series, of multiple substantial uncertainties. The result is robust in the sense that no particular one of these uncertainties (and no small subset, if the suggested smaller points of departure are adopted) is essential for reaching this conclusion.

However, the conclusion that the 1990 CAAA human health benefits are very probably substantially smaller than compliance costs offers a sharp contrast to EPA's more reassuring conclusion that *"**The extent to which estimated benefits exceed estimated costs** and an in-depth analysis of uncertainties **indicate that it is extremely unlikely the costs of 1990 Clean Air Act Amendment programs would exceed their benefits under any reasonable combination of alternative assumptions or methods identified during this study.**"* What explains the difference? The key is that *EPA's uncertainty analysis ignores discrete uncertainties,* such as those we have enumerated, in favor of a continuous, non-negative probability distribution – the Weibull distribution – that puts zero probability density on zero or negative numbers as possible values. Applied to the crucial C-R coefficient, the Weibull uncertainty distribution implies 100% confidence that the coefficient is positive, notwithstanding the numerous studies that have reported zero or negative coefficients. This framing of the uncertainty analysis only allows one to ask (and experts to answer) *how large* are the assumed benefits, but not to question (or provide opinions on) whether they exist. Thus, the major discrete uncertainty about whether positive benefits exist at all – the main focus of our analysis – is simply assumed away in EPA's Weibull uncertainty distribution.

In this context, EPA's conclusion that "it is extremely unlikely the costs of 1990 Clean Air Act Amendment programs would exceed their benefits under any reasonable combination of alternative assumptions or methods identified" is no surprise. It is little more than an assertion of a tautology, similar to, "Assuming that I am right, it is extremely unlikely that any reasonable combination of alternative assumptions would show that I am wrong." But this style of uncertainty analysis fails to cast any light on

uncertainty about the major premise. The not-implausible numbers offered above suggest that that probability of zero human health benefits might well exceed 1 - (0.5*0.5*0.5) = 87.5% based solely on uncertainties about whether there is a positive, causal (not due to confounding or coincidence of trends), non-threshold C-R relation for current and future ambient pollutant concentrations (assuming, conservatively, that there is a 50% probability for each of these three uncertainties – if smaller points of departure are used, then the probability of zero human health benefits exceeds 99%). Thus, assuming that the probability of zero (or negative) human health benefits is zero substantially understates the true uncertainty, and hence substantially overstates the uncertain benefits of the CAAA.

The use of a Weibull uncertainty distribution also conceals other discrete uncertainties, arising from blinkered analysis and failure of imagination. For example, suppose that the CAAA creates more clear days and bright skies. How, if at all, would this affect peak temperatures during the summer, and mortality rates among people without air conditioners? How, if it all, would it affect the future burden of skin cancer? The cost-benefit analysis does not say, and the Weibull uncertainty analysis implicitly disregards all such uncertainties due to drivers of possible costs and benefits not already identified in the analysis. Or suppose that, contrary to our expectations, the CAAA does prolong the last year of life for a substantial fraction of elderly patients suffering from heart, lung, and circulatory diseases. How would this affect health care costs and living costs, and the benefits of longer life, as evaluated by patients and their families? These costs and uncertainties are likewise implicitly assumed away in EPA's analysis, but considering them might further increase the probability that the net benefits

of the CAAA are negative, and further increase the estimated ratio of costs to benefits, beyond the roughly 65-to-1 value (or more) suggested by some of our calculations.

6. Conclusions

EPA's cost-benefit assessment indicates that the CAAA offers the United States an admirable return on investment: an expenditure of only $65 billion per year secures health benefits valued at nearly two trillion dollars per year. If this is credible, any policy maker would be foolish to refuse it. EPA's uncertainty analysis suggests that the large excess of estimated costs over estimated benefits is not only credible, but virtually certain: it is deemed to be "extremely unlikely" that any reasonable changes in assumptions would reverse the sign of net benefits.

However, these promising conclusions depend essentially on the use of a Weibull distribution in the uncertainty analysis, which tacitly assumes away any possibility of serious, discrete errors or uncertainties in the key assumptions – especially about whether future reductions in $PM_{2.5}$ and ozone will necessarily cause future increases in life expectancy. Yet, this is exactly the uncertainty that a large, data-driven literature questions. Numerous reported findings of unexplained negative and highly heterogeneous positive C-R coefficients call into question the assumption that any such single, positive, causal coefficient exists.

Taking into account such discrete uncertainties suggests that the probability of achieving positive benefits could well be small (e.g., less than 12.5%, or less than 1%, if some BMA results[9] are used) so that EPA's uncertainty analysis assuming that this

probability is 100% may be more misleading than informative. A policy maker to whom the CAAA costs and benefits are presented as: “Pay $65 billion for a chance of probably much less than 20% to achieve benefits of about $300 billion,” as our numbers suggest, might well decide that the cost is too great for the uncertain benefits. Even if all of our suggested specific numbers are rejected and replaced by more carefully developed ones, it seems that considering discrete uncertainties has important effects on the analysis, and that such uncertainties should be presented to policy makers if they are to make well-informed decisions to serve the public interest.

We do not believe that the calculations presented here prove beyond doubt that the CAAA costs far more than it produces in benefits. We do think that they make this possibility very likely, and hence well worth including, instead of assuming away, in any uncertainty analysis of CAAA cost-benefit comparisons.

REFERENCES

1. EPA, 2011. The Benefits and Costs of the Clean Air Act from 1990 to 2020: Summary Report. EPA. March, 2011

2. EPA, 2011. The Benefits and Costs of the Clean Air Act from 1990 to 2020

3. Yabroff KR, Bradley CJ, Mariotto AB, Brown ML, F Estimates and projections of value of life lost from cancer deaths in the United States. J Natl Cancer Inst. 2008 Dec 17;100(24):1755-62

4. Aldy JE, Viscusi WK. Age differences in the value of statistical life: Revealed preference evidence. Review of Environmental Economics and Policy. 2007. 1(2):241-260.

5. Murphy K, Topel R. The economic value of medical research. In Murphy K and Topel R. (Eds), *Measuring the Gains from Economic Research: An Economic Approach.* University of Chicago Press. 2003: 9-40. http://faculty.chicagobooth.edu/kevin.murphy/research/murphy&topel.pdf

6. Cox LA Jr. An exposure-response threshold for lung diseases and lung cancer caused by crystalline silica. *Risk Analysis.* 2011

7. Stoeger T, Reinhard C, Takenaka S, Schroeppel A, Karg E, Ritter B, Heyder J, Schulz H. Instillation of six different ultrafine carbon particles indicates a surface area threshold dose for acute lung inflammation in mice. Environ Health Perspect. 2006 Mar;114(3):328-33.

8. Clyde M . Model uncertainty and health effect studies for particulate matter. Environmetrics. 2000 Nov-Dec; 11(6):745-63

9. Koop G, Tole L. Measuring the health effects of air pollution: To what extent can we really say that people are dying from bad air? Journal of Environmental Economics and Management. 2004. 47:30-54. See also: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.164.6048

10. Koop G, Tole L. An investigation of thresholds in air pollution–mortality effects. Environmental Modelling & Software 2006. 21(12):1662-73.

11. Thomas DC, Jerrett M, Kuenzli N, Louis TA, Dominici F, Zeger S, Schwartz J, Burnett RT, Krewski D, Bates D. Bayesian model averaging in time-series studies of air pollution and mortality. J Toxicol Environ Health A. 2007 Feb 1;70(3-4):311-5

12. Schwartz J, Coull B, Laden F, Ryan L. The effect of dose and timing of dose on the association between airborne particles and survival. Environ Health Perspect. 2008 Jan;116(1):64-9.

13. Franklin M, Zeka A, Schwartz J. Association between PM2.5 and all-cause and specific-cause mortality in 27 US communities. J Expo Sci Environ Epidemiol. 2007 May;17(3):279-87.

14. Faes L, Nollo G, Chon KH. Assessment of Granger causality by nonlinear model identification: application to short-term cardiovascular variability. Ann Biomed Eng. 2008 Mar;36(3):381-95.

15. Daniels MJ, Dominici F, Samet JM, Zeger SL. Estimating particulate matter-mortality dose-response curves and threshold levels: an analysis of daily time-series for the 20 largest US cities. Am J Epidemiol. 2000 Sep 1;152(5):397-406.

16. Roberts S, Martin MA. The question of nonlinearity in the dose-response relation between particulate matter air pollution and mortality: can Akaike's Information Criterion be trusted to take the right turn? Am J Epidemiol. 2006 Dec 15;164(12):1242-50.

17. Lianne Sheppard, Richard T. Burnett, Adam A. Szpiro, Sun-Young Kim, Michael Jerrett, C Arden Pope and Bert Brunekreef. Confounding and exposure measurement error in air pollution. Air Quality, Atmosphere & Health 2011.

18. Green LC, Armstrong SR. Particulate matter in ambient air and mortality: toxicologic perspectives. Regul Toxicol Pharmacol. 2003 Dec;38(3):32

19. Gold M, Siegel JE, Russell LB, Weinstein MC (Eds), *Cost-Effectiveness in Health and Medicine.* Oxford University Press. New York. 1996. www.amazon.com/Cost-Effectiveness-Health-Medicine-Marthe-Gold/dp/0195108248.

20. Simkhovich BZ, Kleinman MT, Kloner RA.Particulate air pollution and coronary heart disease. Curr Opin Cardiol. 2009 Nov;24(6):604-9.

21. Simkhovich BZ, Kleinman MT, Kloner RA. Air pollution and cardiovascular injury epidemiology, toxicology, and mechanisms.J Am Coll Cardiol. 2008 Aug 26;52(9):719-26.

22. Adabag AS, Luepker RV, Roger VL, Gersh BJ. Sudden cardiac death: epidemiology and risk factors. *Nature Reviews Cardiology* 2010 (Apr) **7**, 216-225

23. Capewell S, Ford ES, Croft JB, Critchley JA, Greenlund KJ, Labarthe DR. Cardiodiovascular risk factor trends and potential for reducing coronary heart disease mortality in the United States of America. Bulletin of the World Health Organization. 2010 (Feb) 88(2):81-160. www.who.int/bulletin/volumes/88/2/08-057885/en/index.html (Last accessed 7-17-2011).

24. Krewski, D., M. Jerrett, R.T. Burnett, et al. (2009). Extended follow-up and spatial analysis of the American Cancer Society study linking particulate air pollution and mortality. Res Rep Health Eff Inst. 140: 5-114