DRAFT FINAL REPORT

(SAMI did not finalize this report)

MORTALITY RISK REDUCTIONS AND ECONOMIC BENEFITS OF ALTERNATIVE SAMI AIR QUALITY STRATEGIES

Prepared by:

Mathtech. Inc.
202 Carnegie Center, Suite 111
Princeton, NJ 08540

Prepared for:

Southern Appalachian Mountains Initiative (SAMI)
59 Woodfin Place
Asheville, NC 28801

December. 2001

Table of Contents

INTRODUCTION

This report describes the potential mortality risk reductions and economic benefits of air quality improvement actions defined by the Southern Appalachian Mountains Initiative (SAMI). This analysis is part of a broader Integrated Assessment that addresses the steps needed to complete a full examination of incremental air quality improvements. The major elements of the Integrated Assessment include atmospheric modeling, emission inventory development, environmental effects, socioeconomics and policy recommendations. This document reports one element of the socioeconomics area, human mortality associated with exposure to particulate matter less than 2.5 microns in aerodynamic diameter ($PM_{2.5}$). Other topics covered in the socioeconomics portion of the Integrated Assessment include competitiveness, lifestyle changes, visual range impacts, fishing and sense of place concerns. These last five topics are discussed in separate documents being prepared for SAMI.

The six socioeconomics topics covered for SAMI are narrow in scope relative to the possible set of effects. Research conducted by and for the U.S. Environmental Protection Agency (EPA) has identified a variety of outcomes that may be affected in a positive or negative manner with variations in ambient $PM_{2.5}$ (USEPA, 1996). Nevertheless, prior socioeconomic studies indicate that changes in human mortality risk can be a major factor in decisions based on a benefit-cost analysis framework. As a result, while the findings in this report represent only one piece of the information base needed to complete a formal analysis of benefits and costs, they do provide a meaningful part of the total information package.

Purpose

The purpose of this analysis is to provide SAMI with quantitative and qualitative estimates of mortality risk reductions for alternative air quality improvement actions. This task is consistent with the SAMI mission statement:

> "SAMI's Mission: Through cooperative effort, identify and recommend reasonable measures to remedy existing and to prevent future adverse effects from human induced air pollution on the air-quality related values of the Southern Appalachian Mountains, primarily those of Class I areas, weighing the environmental and socioeconomic implications of any recommendations." (SAMI website, 2001)

SAMI research efforts emphasize protection of Class I areas. These are areas classified by EPA as unique due to their natural state (See Figure 1). However, air quality improvement actions will likely impact both Class I areas as well as parts of the SAMI region that fall outside of Class I areas. For this reason, mortality risks, which are population-based, are expected to be an important aspect of any action designed to maintain or improve air quality in Class I areas. Our objective is to rely on prior studies of mortality risk reduction for air quality improvements to determine possible impacts in the SAMI region.[1] These estimates will be available to SAMI decision-makers to assist them in current and future decisions concerning air quality.

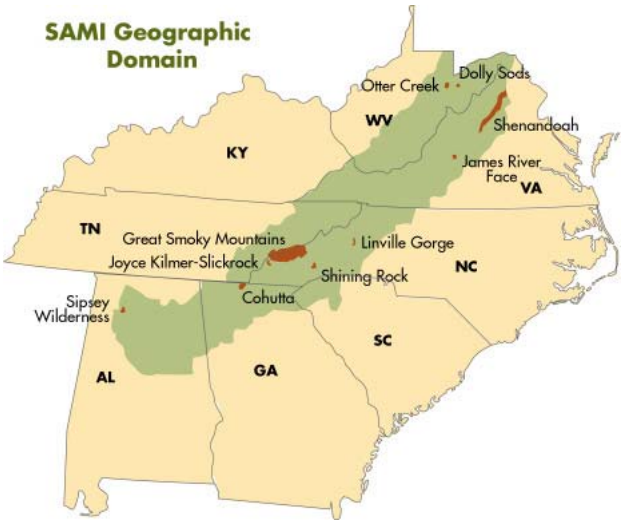


Figure 1: SAMI Domain and Class I Areas

Approach

The analysis follows the outline shown in Figure 2. The concept of an *Integrated Assessment* is made operational by including input data from the atmospheric modeling and emission inventory components in the socioeconomic analysis. These analyses, performed under separate contracts, provide $PM_{2.5}$ air quality concentrations and current/projected population data,

[1] The SAMI region includes eight States in the southeastern United States: Virginia, West Virginia, Kentucky, Tennessee, North Carolina, South Carolina, Georgia and Alabama. The study area for this analysis includes about 85% of the SAMI land area. Coverage is limited by available modeled $PM_{2.5}$ concentrations.

respectively. These data are available for specific locations, years and improvement strategies. The data from these two components of the Integrated Assessment are supplemented with additional demographic data needed to evaluate mortality risks. These input data are then included within statistical models that are based on epidemiological methods that stress the assessment of risk.[2] The evaluation of epidemiological models provides an estimate of "statistical lives saved" for a given change in air quality. This is a key intermediate result of the study and provides SAMI with one indicator of the relative impact of alternative actions.

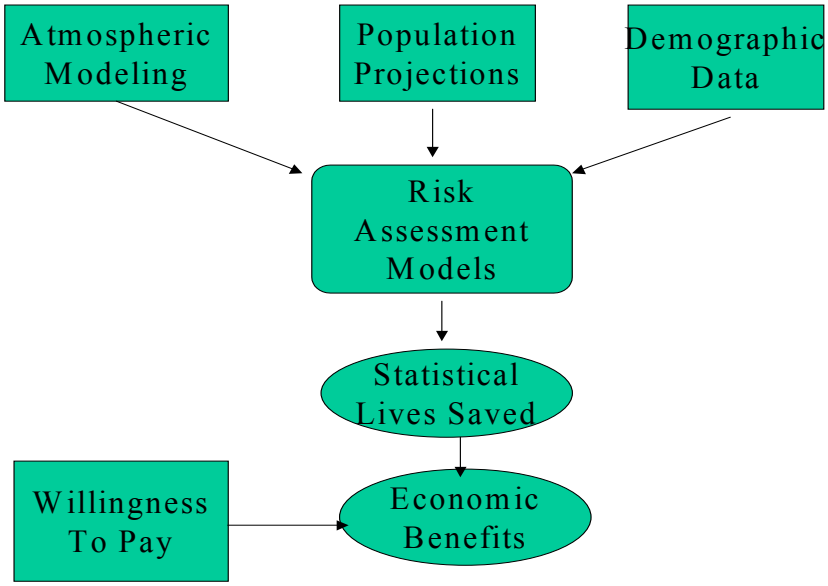


Figure 2: Flowchart of Approach

The transformation of statistical lives saved into an equivalent economic value (in dollars) allows direct comparison of the gains experienced by society (in terms of lives saved) versus the costs to society of the proposed air quality improvement actions. This step is shown at the bottom of Figure 2. This report provides details about both the risk assessment phase (involving statistical lives saved) and the economics phase.

[2] This analysis is driven by the results of epidemiological analyses, a statistical method. Other approaches that help explain the association between mortality risk and exposure to air pollution include toxicology studies and surveys of expert opinions. Epidemiological relationships are a standard way to characterize the statistical associations of interest.

A concern throughout this document is the uncertainty associated with the data and methods. Assessment and valuation of mortality risks is controversial. Concerns span the range of the statistical propriety of estimated models, data reliability and interpretation of the possible causal properties of the analysis. A variety of assumptions are made to conduct the analysis. Although these assumptions are made explicit, each reader will need to judge the relative significance of assumptions and the potential impact of the assumptions on the results. We do not present results for a series of "options" that combine assumptions in a variety of ways. The scope of results that would need to be assessed with such an approach quickly becomes unmanageable. Instead, we present a core set of quantitative results and supplement these core findings with further discussion of "sensitivity" issues.

Summary Findings

Table 1 contains summary data on the air quality concentrations modeled under various SAMI strategies. The table reports the annual average of $PM_{2.5}$ for an historic (observed) situation, three modeled strategies and two years. Mean values are computed at the county level and area-weighted State means are derived to obtain the table entries. SAMI's atmospheric modeling team developed the concentration data.[3] The three strategies are On-the-Way (OTW), referenced as A2, which is used as a comparison baseline, and two control strategies, Bold with Constraints (BWC or B1) and Beyond Bold (BB or B3). These strategies reflect the adoption of alternative emission control strategies for polluting sources (citation). B3 is more restrictive than A2 and B1. B1 is more restrictive than A2. The strategy definitions are based on a co-operative effort between the atmospheric modelers and the emission inventory contractors. Data are provided for two horizon years: 2010 and 2040. We report summary results for these two years, although the economic calculations are performed on a year-by-year basis. Additional discussion of these table values is provided in a later section.

[3] The modeling team has determined that modeled concentration values under-report actual concentrations when compared against observed values. The variation increases with increasing concentration values. See (citation) for additional detail.

Table 1: Summary of Atmospheric Modeling Data:

Annual Mean $PM_{2.5}$ Concentrations in $\mu g/m^3$

	BASELINE 1991-95	2010 A2	2010 B1	2010 B3	2040 A2	2040 B1	2040 B3
Alabama	12.57	12.92	11.95	9.31	12.49	10.50	8.67
Georgia	12.45	12.81	11.68	8.29	12.52	10.26	7.80
Kentucky	10.35	10.50	9.99	8.06	10.04	9.21	7.82
North Carolina	11.47	11.73	10.46	7.06	10.81	8.47	6.51
South Carolina	11.19	11.44	10.43	7.00	10.58	8.73	6.47
Tennessee	12.26	12.25	11.08	8.03	11.60	9.22	7.33
Virginia	11.30	10.90	10.08	7.92	10.41	8.57	7.14
West Virginia	9.86	9.76	9.34	8.13	9.40	8.53	7.58
Avg. in Modeled Area	11.38	11.47	10.58	8.00	10.94	9.20	7.46

Table 2 reports the mean $PM_{2.5}$ concentration changes by State and strategy. The entries are the difference in values reported in Table 1 for a strategy definition. It is the change in concentrations that is the driver behind the predicted change in mortality risk.

Table 2: Summary of Mean $PM_{2.5}$ Concentration Changes in $\mu g/m^3$
By State and Strategy

	2010 (A2-B1)	2010 (A2-B3)	2040 (A2-B1)	2040 (A2-B3)
Alabama	0.97	3.61	1.99	3.82
Georgia	1.13	4.52	2.26	4.72
Kentucky	0.51	2.44	0.83	2.22
North Carolina	1.27	4.67	2.34	4.3
South Carolina	1.01	4.44	1.85	4.11
Tennessee	1.17	4.22	2.38	4.27
Virginia	0.82	2.98	1.84	3.27
West Virginia	0.42	1.63	0.87	1.82
Avg. in Modeled Area	0.89	3.47	1.74	3.48

Table 3 reports summary findings on statistical lives saved. The entries are differentiated by model findings from epidemiological studies and by strategy/year combinations. The new element in Table 3 is the introduction of selected models for our quantitative analysis. As described in a later section, three models have been identified as "core" for this analysis. These models are obtained from the Krewski et al. (2000) re-analyses of Dockery et al. (1992) and Pope et al. (1995) and Lipfert et al. (2000). The entries in Table 3 represent the point estimate for each model for each combination of year and air quality improvement action. Variations across

models in Table 3 reflect differences in estimates of excess mortality risk associated with each model. For any single model, differences in impacts across States (not shown) occur because of variations in population densities and the spatial distribution of air quality improvements. A later section presents maps of the SAMI region that highlight the spatial variability of the statistical lives saved measure.

Table 3: Estimated Statistical Lives Saved by Model and Strategy*

	2010 (A2-B1)	2010 (A2-B3)	2040 (A2-B1)	2040 (A2-B3)
Krewski/Pope	1,002	3,734	2,598	4,905
Krewski/Dockery	1,130	4,212	2,932	5,535
Lipfert	-54	-207	-143	-273

* The evaluation of each model is based on applicability criteria appropriate for that model. See the Models section for more detail.

Table 4 reports the last key element of this analysis. The table shows annualized benefits of alternative strategies by model. The main input data for the transformation from statistical lives saved to economic benefits is an estimate of the value for a statistical life. The development of this measure is described in the Economic Valuation and Benefits section. Note, at the direction of SAMI's Socioeconomic Workgroup (SEWG), no pooling across model results is made. The values represent the mortality risk reduction benefits for a stream of benefits occurring between 2010 and 2040. The economic values are reported as annualized values per million dollars for a statistical life saved. The entries are in 2000 year dollars, discounted to the year 2000. A real discount rate of 5% is used. These economic values are important because they allow direct comparison of the economic benefits to estimated economic costs of the proposed air quality improvement actions.

TABLE 4: Economic Valuation of Annualized Benefits per $1M (2000 $) VSL
by Study and Strategy

Study Characteristics			Strategy	
Study	RR	Range	A2 to B1	A2 to B3
Krewski/Pope	1.06	24.5	$571,185,857	$1,581,344,432
	1.12		1,108,340,946	3,059,081,091
	1.19		1,697,091,178	4,667,080,450
Krewski/Dockery	1.09	18.6	1,250,351,340	3,451,222,219
	1.28		3,550,898,545	9,682,874,209
	1.49		5,688,222,810	15,340,654,963
Lipfert	0.9119	15.09	-60,329,233	-169,863,451

Individual sections of the report provide detail on the data, assumptions and methods that go into producing the quantitative findings summarized in Tables 1 to 4. The tables are presented here to provide initial information about the possible scale of mortality effects in the SAMI region. However, the full report should be read to gain an understanding of the complete scope and limitations of these results.

Organization of the Report

Following this Introduction, the report is organized into three sections: Background, Risk Assessment, and Economic Valuation and Benefit Analysis. The last three sections discuss models, data, remaining issues and results. In cases where detailed tables are informative, selected results are included in the body of the report.

BACKGROUND

Socioeconomic impacts of air quality improvement strategies in the SAMI region were first addressed in 1998 in a Recommendation Phase. At that time, a literature review was conducted to evaluate and rank various endpoints. (cite Phase I Final Report) The effects of concern included air quality related values (e.g., aesthetic values), impacts on materials, agricultural and forestry effects, human mortality and morbidity risks and the economic impacts of emission management actions. These areas were evaluated in terms of expected impact magnitude, quality of the estimate, distribution differences among demographic groups and the

anticipated cost to perform a socioeconomic analysis of the effect. A series of workshops with SAMI's Socioeconomic Work Group (SEWG) narrowed the recommended set of effects to be evaluated to visibility, fishing and community sense of place (each an air quality related value), mortality risk reduction, competitiveness effects on industry (producers) and lifestyle (behavioral) effects on consumers.

The present report represents one part of Phase II of the socioeconomic component of SAMI's Integrated Assessment. We describe and estimate the effect of air quality improvements on mortality risk reductions and the economic value of those risk reductions. The scope of this analysis is determined both by SEWG consensus reached during Phase I and our technical proposal. Two decisions limit the scope. First, the analysis is limited to mortality risk reductions. The analysis does not consider a variety of morbidity risks that may be associated with changes in air quality. Second, the analysis examines the association between mortality risk and fine particles represented by $PM_{2.5}$. Other air pollutants have been associated with excess mortality risk, but changes in ambient concentrations of these other pollutants are not directly evaluated for their potential effect on mortality risk.

The SAMI mission statement quoted earlier emphasizes "cooperative effort" among the various stakeholders who participate in the SAMI process. This design allows for spirited discussion of methodology, acceptable assumptions and presentation style. It also means that internal consensus about an issue becomes a key decision point for project direction. One such decision was made by the SEWG early in this project. The decision involved an agreement on the set of studies to be considered in the mortality analysis. There are several analytical frameworks that have been applied to examine mortality risk. Variations in analytical design respond to whether chronic or acute exposures are emphasized, the temporal and spatial dimensions of the data and the level of aggregation of data. The SEWG recommended that this analysis of mortality risk focus on studies that utilize the prospective cohort design. These studies are generally interpreted to represent both short-term as well as long-term responses to air quality changes. The recommended studies are identified in the next section and a description is provided of data characteristics and methodology. While other published research could contribute to the analysis of mortality risk, the recommendation of studies by the SEWG defines the acceptable scope for the present analysis.

This analysis adopts a "benefit transfer" approach for evaluating mortality risk reductions. This means that no new research is conducted to determine the quantitative associations between exposure and risk. Rather, we rely on previously developed relationships, in combination with data specific to the SAMI region, to infer mortality risk changes. This process makes use of the assumption that results derived in other settings and for other time periods can be exported to a SAMI scenario that covers the period 2010 to 2040. We emphasize that this analysis does not simply re-scale prior risk and benefit findings to better represent the SAMI situation. The analysis follows the evaluation approaches adopted by original researchers. The transfer aspect of our study involves the use of certain key results from the earlier studies as part of the evaluation process.

RISK ASSESSMENT

Risk assessment is a formal procedure for quantifying relationships between a defined risk and a factor that is suspected of altering risk. The steps in a risk assessment involve risk identification, risk characterization, measurement of a risk relationship and risk management. Our emphasis in this section is on analyses that measure risk relationships. We rely on a statistical approach, characterized as epidemiologic, to identify the quantitative association between mortality risk and exposure to $PM_{2.5}$. By itself, epidemiologic information is typically not sufficient to infer a cause and effect relationship between variables in the statistical models. Generally, additional supporting data from other types of analyses, such as toxicology experiments or controlled laboratory designs, need to be considered to support any causal inferences. While there is an expanding body of literature that considers mortality risk from other than an epidemiologic perspective, we do not reference or review that literature in this study. Therefore, the results reported below are best viewed as indicative of the degree of possible association between the risk measure (mortality) and the initiating factor ($PM_{2.5}$ exposure) .

Approaches

Within the body of epidemiologic studies, analysts have used a variety of analytical formats. Oftentimes, the decision about format is more data-driven than influenced by methodological concerns. Two general formats that have been used are time series and cross-section. A time series study, as the name suggests, examines risk relationships over time. Data

on the risk event, the initiator (e.g., air pollution exposure) and other factors that might explain variations in the risk event are collected for multiple time periods and statistical methods are used to identify underlying associations. A key concern is whether unmeasured variables may change over time and represent the "true" causal agent. A cross-section study evaluates changes across the observation units at a given point in time. For example, using data collected for many cities, average mortality rates in a city might be associated with average pollution concentrations. Variations in both mortality rates and concentrations across cities can provide a basis for estimating a statistical relationship. A possible limitation of the time series and cross section study designs is that these formats often rely on aggregate data to represent the location or time period. For example, mortality rates might be defined as the number of observed mortality events in a city, standardized by a specific population level. Aggregate data, also referred to as "ecologic" data in the literature, will often result in less variation (and hence explanatory power) in a statistical analysis relative to data on individuals.[4]

In response to recognized limitations in time series and cross section studies, recent analyses have adopted an analytic framework that is referred to as a "prospective cohort" design. In this design, a group ("cohort") is identified at a point in time and the vital status (mortality) of individuals in the cohort is recorded for a specified follow-up period.[5] At the initial enrollment of the cohort, other individual data items are also collected. The design is prospective in that members of the cohort are selected prior to any knowledge of when they will die. The use of cohorts allows for individual identification of personal traits that may be important in the explanation of variations in mortality risk. Smoking status of an individual cohort member is such a variable. The prospective cohort studies that have been completed are designed to allow for spatial variation in the data and to represent large sample sizes. Both of these factors help increase the precision of the statistical estimates.

To date, the prospective cohort studies have been unable to acquire personal exposure measures for individual cohort members. As a result, community-wide average ambient concentrations collected from fixed site monitors, have characterized personal exposure. This can

[4] Aggregated variables need not be all bad. One might assume that some city-level variables provide context for the risk relationship of interest. For example, average family income might (a hypothesis) be negatively associated with mortality rates. In general, such variables are really serving as proxies for a more complex behavioral action that drives the statistical association. From a risk management standpoint, it would be better to model the behavioral responses than to rely only on the proxy variable.

[5] Cohort members who survive through the follow-up period are censored at the end of the sample period.

lead to measurement error problems for the exposure variable ($PM_{2.5}$) and affect the statistical interpretation of results. The units of the pollution variable also come into play in another way. Concentrations can be recorded in various temporal dimensions. Common statistics are 24-hour averages, annual means or medians and extreme values of the distribution of hourly concentrations. The various ways exposure can be defined plays a role in the interpretation of the type of mortality risk that is being measured. Variations in short-term averages are likely related to variations in acute mortality events. Similarly, long-term averages may provide information about chronic mortality. In this study, we limit our focus to long-term measures of pollution exposure.

In the next section, we discuss the three studies recommended by the SEWG. Each study uses the prospective cohort design. Statistical modeling of these data is based on the Cox Proportional Hazard Model. This model is exponential in form and identifies variations in base survival probabilities as a function of risk factors, such as pollution exposure. Although the basic form of the Cox Model is well defined, variations in how data are used and the impact of specification issues provide a fertile ground for exploratory analysis.

Studies

The studies are Krewski et al. (2000) and Lipfert et al. (2000). The former study is a re-analysis of two earlier prospective cohort studies (Dockery et al. (1992) and Pope et al. (1995). The Krewski re-analyses involve both data validation and new analysis. The two earlier studies played key roles in EPA's efforts to establish a new Fine Particulate ambient air quality standard in 1997. The Lipfert study is relatively new and additional work is in progress. We emphasize that among these studies, there are effectively three independent population cohorts. A confirmation of findings between, say, the original Dockery study and the Krewski re-analysis of Dockery should be interpreted as a confirmation of numerical methods, not a new data point about excess mortality risk. In this section, we describe the attributes of the cohort samples and discuss general features of each cohort.

Dockery et al. (1992) -- This study is part of the well-known Harvard Six Cities Study. Cohort members were white adults, aged 25 to 74, living in Watertown, Massachusetts, Steubenville, Ohio, Harriman, Tennessee, St. Louis, Missouri, Portage, Wisconsin or Topeka, Kansas. The recruitment of individuals for the study was done in a random manner, with

reasonable response rates cited by the authors. The study was specifically designed to estimate the effects of air pollution on various causes of mortality. Individual control factors were collected for each cohort member. These data included questions about age, gender, weight, height, education, smoking history, potential for occupational exposure and existing medical status. A summary of selected sample characteristics is reported in Table 5.

Data were initially obtained in 1974 in Watertown, with cohort members in other cities first enrolled up to 1977. The period of follow-up for the cohort ranged from 14 to 16 years, with study termination in 1991. A noteworthy feature of the study is the assessment of individual data at selected intervals following initial enrollment. Although Dockery and his co-authors did not report findings with this additional information, subsequent researchers (Krewski et al., 2000) have used the information to advantage.

The individual data on cohort members was supplemented with ambient concentration data collected at sites operated by the study sponsors. The sites are described as centrally located in each of the six cities. It does not appear that monitored data from sites not specific to the study were used to compute a city-wide average value. The use of monitors located and operated for study purposes provides a level of control over the data. However, there is some loss in variation in monitored data that is often observed across an urban area. $PM_{2.5}$ data were collected from 1979 through 1985. Other pollutants were monitored, but later statistical analysis found the most robust results with $PM_{2.5}$ included as an explanatory variable for mortality rate. No multi-pollutant specifications were reported. The observation-weighted average of annual mean concentrations across the six cities is 17.8 $\mu g/m^3$. The individual city mean values range from 11.0 $\mu g/m^3$ in Portage to 29.6 $\mu g/m^3$ in Steubenville. This range is used to characterize the standardized relative risk values reported below for individual model results.

Table 5

Selected Sample Characteristics by Cohort

Factor/Cohort	Dockery	Pope	Lipfert
Sample size	8,111	295,223	26,721 (for $PM_{2.5}$)
Years of Follow-up	13.7	7	Approx, 9
Age (yrs)	49.6	56.6	51.2
Gender (% Female)	54.7	55.9	0
Ethnicity (% White)	100	94	64.7 (non-Black)
Smokers (%)	36	51	56.7
Not HS grad (%)	28.1	11.3	N.A. (individual)

Pope et al. (1995) -- Pope is also a prospective cohort study. The source of cohort data was a sample collected as part of the American Cancer Society (ACS) Cancer Prevention Study II. Although the cohort data were not collected with air pollution effects in mind, the type of information collected was appropriate for conducting an environmental risk study. Two cohorts were defined corresponding to the availability of monitored air pollution data. In this report, we focus on the fine particulate cohort ($PM_{2.5}$) which included 295,223 cohort members residing in 50 U.S. cities at enrollment.

The Pope cohort included individuals first enrolled in the study in 1982, with follow-up through calendar year 1989. Reviewers of Pope have noted that sampling of the population was not random, so that generalization of the study findings must be done with care. Cohort members were restricted to adults 30 or older living in a household with at least one member 45 years old or more. Individual data were collected for age, gender, weight, height, demographic characteristics, smoking history, alcohol consumption, occupational exposure and medical status. An independent reviewer reported the cause of death. Table 5 includes a summary of selected sample characteristics for the Pope cohort.

The Pope study extends the scope of prior mortality risk studies to a larger geographic setting. The availability of individual data, which characterizes the prospective cohort design, is retained and wider geographic variability (50 cities) is introduced. This broader design allows for additional hypothesis testing. The limiting factor in the geographic scope is the availability of ambient monitored data. Pope includes all monitored data meeting minimal reporting criteria and computes averages of monitored data across locations in a city with multiple monitors. The median of the annual $PM_{2.5}$ distribution is reported and used in subsequent statistical analysis. This is a different statistic than the mean, which is more commonly used in other studies. The measure of $PM_{2.5}$ exposure for a cohort member is represented by the median in the year just prior to entry into the cohort.

Lipfert et al. (2000) -- A cohort of U.S. male military veterans who participated in a Hypertension Screening Program in the mid-1970's formed the observation set for this study. Individual characteristics recorded at entry into the screening program included age, race, systolic and diastolic blood pressure, smoking history and medical history. The Lipfert study considers only all-cause mortality. Table 5 records average values for selected characteristics of the Lipfert

sample. These entries appear to pertain to the larger cohort of veterans in the screening program rather than members matched with a particular pollution measure.

Lipfert included a series of "ecologic" variables to represent the contextual setting within which a cohort member resided. Examples of the ecologic variables are income, socioeconomic measures and environmental measures such as altitude and humidity. The in-depth evaluation of these ecologic variables is one feature that distinguishes Lipfert from the other cohort studies.

The Lipfert analysis also differs from other prospective cohort studies in the manner in which follow-up time intervals are analyzed. Lipfert segments the cohort into a series of 12 exposure/mortality periods. This allows a more focused assessment of the stability of the relative risk factors over time. To achieve this temporal detail, the length of an analysis period is more restricted than the data might otherwise allow.

$PM_{2.5}$ data were obtained from EPA's Inhalable Particulates Network database and collected for the period 1979-1984. Lipfert defines the spatial unit in terms of the county. A total of 103 counties relate PM exposure to individual and ecologic characteristics.

Models

The prospective cohort databases can be evaluated with a variety of statistical specifications (or models). The studies described above, and re-analyses of two of those studies, have relied on the Cox Proportional Hazard Model as the basic estimation form. This model is exponential and describes a particular mathematical relationship between the probability of death and factors that may explain variations in the probability of dying. In equation form:

$$Y = B*EXP(\blacktriangleright * PM_{2.5})$$

Where: Y is the hazard rate or the probability of dying[6];
B is a function of a vector of variables believed to affect mortality;
$PM_{2.5}$ represents concentrations of fine particles;
$\blacktriangleright$ is a coefficient to be estimated; and
EXP is the exponential function.

[6] Technically speaking it is $-\frac{ds(x)}{dx}$, where $s(x)$ measures the probability of survival to age x.

The statistical analysis involves estimation of different specifications of this equation. Initially, an analyst might retain the basic structural form of the equation but evaluate alternative sets and transformations of potential explanatory factors. Eventually, other structural (mathematical) forms might be estimated to better address specific questions about the relationship of interest.

In this section, we examine results based on the Cox Proportional Hazard Model specification. We identify three models, one model for each cohort. These models are reported in the Krewski reanalysis of Dockery, the Krewski reanalysis of Pope and the Lipfert study. Each of these studies examines a series of possible models to describe the mortality risk relationship.

Krewski/Dockery: The Krewski reanalysis of Dockery (K/D) extended the set of co-variates considered by the original study. An alternative risk model, characterized as "full", was estimated which included more than 40 co-variates. This compares to the original study that identified eight co-variates. The additional variables include data collected from cohort members at enrollment and transformations of previously used variables. Statistical tests were then used to trim the list of co-variates. The 24 variables that remain are referred to as the extended model. We use the extended model of K/D as the core model for this cohort. This selection reflects a judgement that the extended model provides the best use of the available cohort data that accounts for confounding variables and effect modifiers. Other specifications are considered by K/D. However, we believe these alternatives either demonstrate the advantage of the extended model or do not clearly resolve the issue that prompted the alternative specification.

The extended model was stratified by age (five year intervals) and gender to provide individual measures of baseline hazard for each age-gender combination. Co-variates in the extended model include:

- Tobacco Consumption (7 variables)
- Education Level (2 variables)
- Occupational Exposure (1 variable)
- Body Mass Index (2 variables)
- Marital Status (3 variables)
- Alcohol Consumption (3 variables)
- Interaction with Gender (3 Marital Status variables)
- Interaction with Gender (3 Alcohol Type Consumption variables).

Several mortality endpoints were examined in the original study and re-analyses. This analysis examines results associated with all-cause mortality only.

Krewski/Pope: The Krewski reanalysis of Pope (K/P) also extended the set of co-variates considered by the original study. Using an approach like that for K/D, an alternative risk model, characterized as "full", was estimated which included 58 co-variates. This compares to the original study that identified 11 co-variates. The additional variables include data collected from cohort members at enrollment and transformations of previously used variables. Statistical tests were then used to trim the list of co-variates. The 37 variables that remain are referred to as the extended model. We use the extended model of K/P as the core model for this cohort. As with K/D, this selection reflects a judgement that the extended model provides the best use of the available cohort data that accounts for confounding variables and effect modifiers. Other specifications are considered in K/P. However, we believe that these alternatives either demonstrate the advantage of the extended model or do not clearly resolve the issue that prompted the alternative specification.

The extended K/P model was stratified by age (one year intervals) race and gender to provide individual measures of baseline hazard for each age-gender-race combination. Co-variates in the extended model include:

- Tobacco Consumption (14 variables, including exposure to passive smoke)
- Education Level (2 variables)
- Occupational Exposure (1 variable)
- Body Mass Index (2 variables)
- Marital Status (2 variables)
- Alcohol Consumption (4 variables)
- Interaction with Gender (9 Smoking variables)
- Interaction with Gender (3 Occupational exposure variables).

Several mortality endpoints were examined in the original study and re-analyses. This analysis examines results associated with all-cause mortality only. K/P relied on $PM_{2.5}$ data from the original Pope study and supplemented the data with entries from EPA's Inhalable Particulate Network. This extended the city count from 50 to 63 and also provided information to calculate an arithmetic mean statistic for $PM_{2.5}$. The original Pope study conducted their analysis based on a median measure. The recommended core model uses the mean statistic included in K/P.

Lipfert et al. The Lipfert analysis used the Cox Proportional Hazard Model as the primary regression specification. The analysis was conducted in stages to examine incremental changes in risk under various specifications. Preliminary runs were made using the ecological variables collected for the study. Pollution variables were excluded in order to better understand the independent role of the ecological variables. Lipfert also specified a screening model using indicator variables for VA Center location as a proxy for air pollution. This model included over 200 terms, many defined as categorical variables to allow for non-linear responses. The variables in the screening analysis included age (4), systolic blood pressure (7), diastolic blood pressure (7), body mass index (7), age interactions (multiple), race (2), smokers (4), VA Center location (32). Next, a baseline model was estimated with county air quality concentrations in county of residence replacing the hospital indicator variables. A final model added terms for height and 11 ecological variables. Both the baseline and final models identified 10 age categories.

A unique feature of the Lipfert analysis is the definition of a combination of 12 exposure periods and mortality periods. This allowed for specific hypotheses to be tested about temporal variations in estimated excess risk. A single period specification (defined for the full follow-up period) is also presented to permit more direct comparison to study designs of other researchers.

The Lipfert study considered a variety of pollution measures. Results for eight separate single pollutant models are reported. We focus on $PM_{2.5}$. Based on discussion in the paper and guidance from the lead author, the final model with ecological variables is the basis for our core model selection. We further refine the core selection to the $PM_{2.5}$ pollutant with exposure over the period 1982-84 and the mortality event occurring in the 1982-88 mortality period. This is identified as a concurrent exposure/mortality segment.

Summary of Core Results

Table 6 summarizes the relative risk estimates for the core models. The table records the study name, the reference for the core model, the relative risk measure from the selected model, the 95% Confidence Interval for the risk estimate, and the range in $PM_{2.5}$ for which the relative risk is reported. The studies do not report a full set of statistics for all co-variates in the selected specifications. However, given the mathematical form of the Cox Proportional Hazard Model, it is only the coefficient (or the equivalent relative risk measure) that is needed to evaluate risk reductions as a function of $PM_{2.5}$ changes.

Table 6: Recommended Core Models for SAMI Mortality Study

Study	Core Model Reference	Relative Risk Measure	95% Confidence Interval	$PM_{2.5}$ Range
Krewski et al./ Dockery	Table 7	1.28	(1.09 to 1.49)	18.6
Krewski et al./ Pope	Table 31 (mean)	1.12	(1.06 to 1.19)	24.5
Lipfert et al.	Table 7	-0.092*	NA	15.09**

*Fractional risk based on mean value of $PM_{2.5}$ minus estimated background
**Implied relative range based on full sample statistics

SENSITIVITY ISSUES

The core model results shown in Table 6 make use of an extensive set of data and transformations of that data. All results are developed using the basic Cox Proportional Hazard Model. This model relies on several maintained assumptions that may not apply for the three cohort databases. It is important to address departures from the assumptions of the Cox model and departures from standard assumptions of statistical models in general to better understand the results of the statistical risk analysis.

The term "sensitivity" that identifies this section is a euphemism for the existence of real concerns. However, it is the term used by the Krewski team to describe a series of additional analyses they conduct to more fully assess the Dockery and Pope cohort data. Their analyses do not always describe issues of existence or magnitude for the item of concern. But, formal recognition of the issues is a step in the right direction.

An earlier section of this report referred to mortality risk assessment for environmental changes as "controversial". To a large extent, significant elements of controversy have been absent in the presentation to this point. In this section, we identify some of the issues that contribute to uncertainty in statistical risk analysis. Generally, concerns arise for three reasons. There are problems with the data and the definition of variables. There are maintained assumptions of the selected structural model that may not hold. Finally, the analysis may produce findings that fail to meet accepted statistical requirements.

In this section, we identify a variety of issues that might affect the quality of the estimated risk measures. Many topics are touched on only briefly. However, two of the issues are discussed in some detail. Spatial autocorrelation is afforded extended treatment because the issue has been the focus of considerable discussion in the SEWG. Second, we address the definition of, and data availability for, the measure of pollution exposure.

Estimation of an unbiased measure of risk is a difficult task. The art of risk assessment as applied to environmental problems has matured over time. But there is still room for growth. As a result, opinions can vary about the estimation and interpretation of mortality risk. Individuals must weigh the value of results obtained in "conventional" analyses against remaining uncertainties associated with unresolved (unresolvable?) questions. There is likely no single correct conclusion.

This section and earlier sections should be considered together to assist the reader in forming a personal judgement about quantitative risk analysis. In addition, a body of science exists outside the framework of epidemiological analysis. This other evidence, while not covered here, provides another basis for decisions.

Topics of Concern -- Brief Synopses

Confounding

A variable is a potential confounding variable if it is expected to affect the value of the dependent variable of the specification (in this case, mortality) and is correlated with another explanatory variable included in the specification. As a result of these relationships, it is possible for bias to exist. In particular, exclusion of a confounder from the specification may result in the correlated included variable reflecting not only it's own impact on the dependent variable but also the effect of the missing variable. In this case, too great of an effect might be attributed to the included variable. The selected core models for this study are believed to include important confounding variables. However, one can not be sure that all possible relationships have been quantified.

Constant Risk Factor Over Time

The exposure risk measure in the Cox Proportional Hazard Model is fixed over the cohort follow-up period. This is a maintained assumption of the model and implies that risks are not time

dependent. Preliminary tests by the Krewski team suggest that time dependence could be an issue. This concern is also a fundamental aspect of the Lipfert study design. These findings are important because of generally improving concentrations of $PM_{2.5}$ over time and because of the extended time horizon (to 2040) for the SAMI strategies. A recommended replacement model is not presented.

Flexible Risk Models

The issue of flexible risk models is concerned with testing the Cox model assumption of a linear relationship relative to the log-hazard. Using regression techniques, the Krewski team raises the possibility of the existence of non-linearities. The Lipfert study takes this issue a step farther and examines the possibility of thresholds. However, Lipfert only observes a threshold (a no effect concentration range) when ozone is included as the pollutant explanatory variable. Given the impact of a threshold on the benefit calculations, additional research will be important for this question. The EPA has noted that evidence for mortality effects below 15 $\mu g/m^3$ concentration level of annual average $PM_{2.5}$ is less certain than at higher concentration levels. In part, this guides Agency thinking on the setting of primary health standards. We do not impose a threshold in the calculation of mortality risks in this analysis. One reason is that the hazard model for which relative risks are estimated would be incorrectly evaluated in the range of concentrations above the stated threshold.

Population Mobility

Population mobility involves the possible existence of measurement error in the pollutant variable. If a cohort member moves from the listed place of residence prior to enrollment in the cohort or during the follow-up period, the characterization of exposure at the city of residence at enrollment may not provide an accurate indication of exposure. The original Dockery study included data on length of residence at the enrollment city as well as follow-up questions at 3, 6, and 12 years into the study. The K/D re-analysis examined these data to identify possible mobility effects. Generally, because of the relatively small number of "movers" in the cohort sample, the reanalysis team was not able to draw firm conclusions.

Co-pollutants

There is often a high correlation among various pollutants in the atmosphere in a given location. If a model includes a single pollutant, that variable might reflect negative impacts of other pollutants as well. The Krewski reanalysis finds that SO_2 may play a prominent role in explaining variations in mortality risk with an attendant reduction in the coefficient of $PM_{2.5}$. This appears to be a topic that has yet to be adequately explained and should be the focus of additional research. It is especially relevant in the context of risk management strategy definition.

Ecological co-variates

Previous reviews of the Dockery and Pope studies have expressed concern with the ecologic nature of the pollutant variables. Other variables in the statistical analysis represent data specific to a cohort member. Inclusion of other ecologic variables in the Cox specification has been recommended to better understand the possible existence of confounding at the ecologic level. The Krewski reanalysis examines twenty ecological variables associated with cities in the Pope $PM_{2.5}$ cohort. The results indicate that selected variables could be confounders. However, the analysts also note that the interpretation of statistical results for ecologic variables in the Cox framework must be approached carefully. In particular, they note that there is a downward bias (toward the null hypothesis of no effect) in adjustments to the confounded exposure measure. This observation may be important for the SAMI mortality analysis since the Lipfert study relies heavily on ecological co-variates in the core specification.

Identification of sensitive sub-populations

An examination of the relationship between exposure to $PM_{2.5}$ and mortality should consider the individual characteristics of the sample cohort. These characteristics may be effect modifiers. This means that the characteristic itself influences the predicted mortality risk. In the Krewski re-analyses of Dockery and Pope, a series of tests is conducted to identify possible effect modifiers. In addition, the analysts examine various categorical definitions of population groups and whether identified effect modifiers may also be confounding variables. Educational attainment is determined to be a strong effect modifier. A cohort member with a high school education or less is found to be at greater risk of mortality then a member with more than a high school education. However, additional tests do not indicate a confounding effect between educational attainment

and $PM_{2.5}$ concentrations. Similar analyses are conducted for other sub-populations and confounding issues appear to be absent.

Occupational confounding

An association may exist between mortality risk and exposure to environmental pollutants in the workplace. Such exposure involves dose over an extended period and may involve a virtual soup of substances. The original Dockery and Pope analyses collected data from their cohorts at enrollment on occupation and self-reported exposures. The Krewski re-analysis effort extended this information with additional data development. They defined an ordinal index of dirtiness based on occupational classification and an index of exposure to accepted lung carcinogens. Statistical tests were conducted to assess the possibility of confounding between exposure to workplace pollutants and ambient exposure. The analyses did not find evidence of such confounding.

Accuracy of mortality data

The endpoint of the selected studies is a measure of mortality. In epidemiological work, more robust results may be obtained if the exposure measure is associated with a specific underlying cause of mortality. A specific cause may be suggested by toxicological or physiological evidence. As more specific causes are considered, the possibility of measurement error in the mortality variable should be assessed. In the present analysis, we evaluate only associations that are defined by all-cause mortality (including mortality from non-health related causes). Therefore, concerns about measurement error for the mortality variable should be minimized.

Selection bias

The use of risk equations in a benefit analysis of SAMI air pollution reduction strategies requires that the transfer of risk measures for a cohort can be made in a reasonable fashion. In particular, the characteristics of the cohorts should correspond to the characteristics of the population for which benefits are calculated. This concern has two features. First, does the cohort represent a random sample of populations? Second, what is the scope of characteristics of the cohort? The Krewski reanalysis team notes that the Pope cohort does not represent a random sample of the population. However, no adjustments are suggested. With respect to the scope of

characteristics, the SAMI benefit analysis resticts quantitative calculations to those sub-groups represented in the cohort design. For Dockery and Pope (and their re-analyses) this means that estimates are limited to adults over 25 and 30, respectively. For the Lipfert study, results are reported for a subset of the population that represents male adult veterans with existing mild hypertension.

Topics of Concern -- Expanded Discussions

The last two items discussed as sensitivity issues are spatial autocorrelation and the definition of the $PM_{2.5}$ exposure measure. Spatial autocorrelation involves a departure of the statistical model from accepted assumptions. The latter topic spans issues of measurement error and statistical form. The more detailed discussions allow for a broader treatment of an issue of keen interest to the SEWG (spatial autocorrelation) and a more thorough examination of a series of concerns associated with the measurement of exposure.[7]

SPATIAL AUTOCORRELATION

OVERVIEW

The relationships examined in this report explain variations in mortality rates in terms of a set of explanatory variables. In particular, we are interested in how changes in exposure to $PM_{2.5}$ are associated with changes in mortality rates. This association is a stochastic rather than a deterministic, or exact, relationship. That is, there is random error in the relationship. Spatial autocorrelation refers to a statistical property that is of concern when these relationships are estimated across several different geographic areas.

Researchers often specify stochastic relationships by including an error term in the hypothesized equation. In this case, mortality rates are associated with a set of explanatory variables, including $PM_{2.5}$, and an error term. This error term must have certain properties for the statistical analysis to proceed properly. Among the critical assumptions is that the error terms are independent across observations in the sample of data used to describe the underlying relationship. In the studies considered here, with observations (i.e., members of a cohort) defined for different geographic (spatial) locations, this means that the variables which explain mortality

[7] Critical comments from an SEWG member have been prepared in response to this summary of spatial autocorrelation. Those comments are under review and are not included in the present text.

risk, or mortality risk itself, should not be clustered in groups of higher (or lower) value depending on geographic location. In the standard regression model this means that the error terms should be independent. With spatial autocorrelation this assumption does not hold. As a result, statistical analysis of the data may result in estimators whose statistical power is affected. This impacts judgements about statistical significance. In summary, while spatial autocorrelation reflects certain associations in the data, the problem to be addressed lies in the impact of these associations on the error terms of the model.

The original analysis conducted by Pope et al. (1995) assumes statistical independence of all observations.[8] The reanalysis effort of Krewski et al. both replicates the original analysis under the assumption of independent observations and develops several analytical techniques that recognize the possibility of spatial autocorrelation in the cohort data. For $PM_{2.5}$ exposures, the additional models include a random effects model based on the Cox proportional hazard specification and a series of two-stage models that allow for different levels of spatial modeling. More detailed models, involving filtering techniques, are limited to the sulfate cohort because of the broader geographic scope of that cohort. In the following sections, we describe the alternative models used by Krewski et al. for the $PM_{2.5}$ cohort and present results of the calculations. These results are compared to the results obtained under the independent observation assumption. A final section presents recommendations for the use of these results in the benefit analysis portion of the study.

ALTERNATIVE MODELS

Models based on the Cox Proportional Hazard Function

The reanalysis of the Pope et al. data describes two models that use the Cox Proportional Hazard Model to describe survival (mortality) rates. The first model follows the lead of the original investigators and assumes statistical independence. Cohort variables include individual data and selected ecologic variables. This model is referred to as the "Independent Observations" model. It replicates the original structural model of Pope et al. with an extended set of explanatory variables.

[8] The discussion in this section focuses on the Pope et al. study and the Krewski et al. reanalysis of that study. The reanalysis team effort for the spatial autocorrelation issue was limited to the Pope et al. cohort because of the broader spatial representation of cities in that cohort.

The second model extends the Cox proportional hazard specification to allow the baseline hazard function (defined for specific age groups, gender and ethnicity) to vary at random among cities in the $PM_{2.5}$ cohort. This model is referred to as the Cox Random Effects model. This model is designed to reflect the possibility that members of the cohort who live in the same city are more likely to experience similar mortality risks than cohort members who reside elsewhere. As noted by Krewski et al. this structural design implies that each city has a differential effect on individual mortality risk.

Two-Stage Models

As an alternative to the single equation Cox format, the reanalysis team also develops a framework based on multi-stage regression methods. The purpose of this approach is to eliminate embedded spatial autocorrelation in a first stage model. Then they estimate mortality risk relationships in a second stage regression that utilizes the results of the first stage analysis. In the first stage, only individual level data are used, including dummy (or indicator (0,1)) variables for the cities. The $PM_{2.5}$ data, which is a community (or ecologic) variable, is omitted from the first stage. The inclusion of the city-specific dummy variables plays a role similar to the variation introduced for cities in the Cox Random Effects Model (described above). In the second stage of this approach, the logarithms of the first stage predicted mortality rates are regressed against city specific explanatory factors, including $PM_{2.5}$. The model is designed so that the error terms of the second stage model are statistically independent.

The Cox Random Effects Model and the Two-stage approach are described in Krewski et al. as "Independent Cities Models." This description follows from the prediction of differential mortality rates based on the city of residence of the cohort member.

The Krewski et al. reanalysis team also examines a model which allows spatial autocorrelation to extend to a regional level. This is accomplished in much the same way as the Two-stage city model, but dummy variables are now introduced for seven predefined regions of the United States (the $PM_{2.5}$ cohort encompasses cities across the United States). This structural form allows the analysts to examine whether correlation in the data extends to a regional spatial scale. This model is referred to as the "Regional Adjustment Model." It is included as an alternative two-stage model because the modeling approach utilizes two regression equations to develop independent predictions of mortality risk by region.

Spatial Filtering Models

The reanalysis of the Pope et al. cohort also uses spatial filtering techniques. However, these analyses are limited to the sulfate cohort. This is a result of the more limited geographic scope of the $PM_{2.5}$ cohort, which contains complete data for only 50 distinct cities. Since the present analysis focuses on $PM_{2.5}$ changes, we do not address the spatial filtering models.

APPROACH

The Krewski et al. reanalysis of spatial autocorrelation involves three methods. First, a graphical review of the cohort data is conducted to develop a qualitative understanding of the possible existence of spatial autocorrelation. This method uses maps with overlays of $PM_{2.5}$ concentrations and mortality rates. The second method applies formal statistical tests to determine the existence of spatial autocorrelation. The third method involves estimation of different structural models, as described above. This document looks at the results of the estimation process.

RESULTS

The results of the spatial autocorrelation reanalysis of the Pope et al. cohort for $PM_{2.5}$ are summarized in Table 46 of Krewski et al. A selection of these results is presented in Table 7. For comparison purposes, results for the basic Cox model analysis are also presented.

Table 7: Results of Spatial Autocorrelation Reanalysis for $PM_{2.5}$ Cohort

Model	Relative Risk	Confidence Interval
Basic Cox Model	1.18	(1.09 - 1.26)
Independent Observations	1.18	(1.03 - 1.35)
Independent Cities	1.29	(1.12 - 1.48)
Regional Adjustments	1.16	(0.99 - 1.37)

Source: Table 46 of Krewski et al.

There are several aspects of these results, and the full set of results shown in Table 46, that should be noted. First, the results refer to all cause mortality. Second, it appears that each model reported in Table 46 uses the median of $PM_{2.5}$ as the statistical measure of exposure. The

data we received from the SAMI Atmospheric Modeling team represent the annual arithmetic average of $PM_{2.5}$. When the reanalysis team used the mean of $PM_{2.5}$ in a basic Cox framework, they obtained a relative risk measure of 1.12 (1.06 - 1.19). The risk measure is estimated for the same range in $PM_{2.5}$ as reported for the median analysis, 24.5 micrograms per cubic meter. Finally, Table 46 reports the impact of a series of ecologic co-variates on mortality risk across the various models. In general, these co-variates do not significantly explain variation in mortality rates. An important exception is SO_2.

REVIEW AND INTERPRETATION

Krewski et al. find that spatial autocorrelation is a significant issue for the Pope et al. $PM_{2.5}$ cohort. This conclusion is supported by both graphical analysis and statistical tests. The Krewski et al. reanalysis effort also develops alternative models to address spatial autocorrelation. The results of the reanalysis provide both higher and lower relative risk estimates and somewhat broader confidence intervals. In theory, spatial autocorrelation is expected to affect the interpretation of statistical significance for the explanatory variables. However, the reanalysis also finds a range of risk coefficients. This latter result suggests that an important explanatory variable may have been omitted from the equation. Krewski discusses this possibility and indicates that the supporting graphical analysis might suggest other important explanatory variables. However, no specific new variables are identified in the reanalysis.

The Krewski et al. reanalysis of the $PM_{2.5}$ cohort uses standard methods to identify and correct for the presence of spatial autocorrelation. Similar approaches have been used in the economics literature in cross-section labor studies of wage rates. (Moulton, Brent, "An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Macro Unit," *Review of Economics & Statistics*, May 1990.) Although the reanalysis team uses common statistical methods, their report has not been subject to external peer review at a level that is normally sought.

The spatial autocorrelation analysis uses the median statistic for $PM_{2.5}$ to represent exposure. This is the same measure used in the study conducted by the original investigators. However, it is a different measure than the one available to us from the Atmospheric Modeling component of the SAMI Integrated Assessment. The exposure data we received are annual arithmetic average data. The reanalysis team augments the original data so that a mean value of

exposure can be defined. The reanalysis also estimates the Cox proportional hazard model with these data. But, there is no indication that the mean exposure data were included in the spatial autocorrelation analysis. It would be inappropriate to use risk measures estimated from median estimates of exposure as if they represented mean measures.

RECOMMENDATIONS

We recommend that the spatial autocorrelation analysis be included as a sensitivity result for the SAMI mortality topic area. These results would inform decision-makers about possible ranges in results, but would not be included as a core result of the analysis. We have two reasons to support this decision. First, the spatial autocorrelation analysis has not received the same level of peer review as other aspects of the cohort analysis. There has been independent internal review, but not independent external review that normally is associated with a journal publication. Second, the spatial autocorrelation analysis uses a different measure of exposure than is available for the analysis of SAMI air quality actions.

PARTICULATE MATTER ISSUES

A key variable in the analysis of a possible association between mortality rates and air pollution is the measure of air pollution used by the researcher. In this analysis, we use $PM_{2.5}$ concentrations to explain variations in mortality rates, using statistical methods. It is important that there is a clear understanding of which air pollution variable is being used and that the units of the chosen variable are consistently defined and accurately applied throughout the analysis. To better understand some of the requirements for the exposure variable used in the risk analyses, we examine general issues related to the $PM_{2.5}$ data. The topics covered include: definition, units of measurement, spatial characteristics, temporal characteristics, statistical issues, the form of association with mortality and transfer to the SAMI setting.

WHY CHOOSE $PM_{2.5}$?

Many air pollutants are thought to contribute to excess mortality. Particulate Matter (PM) is an air pollutant that is thought to affect human morbidity and mortality, primarily through impacts on the respiratory system. This hypothesis has been fostered by prior observations of severe PM events in cities and the associated increases in morbidity or mortality. Careful thinking about the physiological potential of PM on the human respiratory system has also supported the possibility of negative health outcomes. Such observations have contributed to PM

being considered one of the principal pollutants (a criteria pollutant) for regulation under U.S. environmental policy. While many factors, including other air pollutants, may contribute to variations in mortality rates, PM is believed to be a key explanatory variable.

SIZE OF PARTICLES

There are many ways to characterize PM. One is by size. Specifically, reference is often made to the aerodynamic diameter of a particle. A particle referred to as $PM_{2.5}$ has a diameter of less than 2.5 microns. This is considered a "fine" particle and is of small enough size to be deposited more deeply in the human respiratory system. Fine particles are often generated from combustion processes. Other, larger particles, can occur as dust in wind-blown material. The history of regulatory concern in the United States has changed over time to reflect different particle sizes. Initially, Total Suspended Particulates (TSP) were the focus of regulation. As more became known about the physiology and toxicology of PM and as measurement and control techniques improved, regulatory focus evolved to PM_{10} and most recently to $PM_{2.5}$.

COMPOSITION OF PM

Another way to characterize PM is by composition. A specific particle may differ from another particle of equivalent size because of differences in the material that constitutes the particle. The relative potential for harm to humans (toxicity of the particle) depends on the elements contained in that particle. Thus, it is very difficult to speak in exact terms about PM. This complicates analysis. For example, an epidemiological analysis that relies on PM data generated from wind blown dust would not be a good predictor for marginal changes in PM generated from combustion sources. This same concern may affect the transfer of relative risk measures between geographic areas. For example, if the elemental components of PM in the SAMI region are expected to be more harmful than the PM included as independent variables in a cross-section analysis, the transferred results may be biased. Studies that examine the "geographic heterogeneity" of risk measures can address this issue.

CONCENTRATIONS VERSUS EXPOSURE

Epidemiological studies typically rely on ambient (monitored) measures of PM.[9] Three concerns arise with data of this type. First, do ambient measures of concentration adequately portray exposure? Second, do ambient measurements fully characterize exposure from ambient and non-ambient sources? Third, do ambient measures adequately reflect individual exposure?

The desired measure of PM in an epidemiological study is the exposure of an individual. Exposure takes into account the deposition of a pollutant in the proximity of an individual as well as the behavioral actions of that individual. A person may take defensive actions to reduce overall exposure. Possible actions might include the allocation of time between indoor and outdoor activities or the use of filtering mechanisms to clean the air in their environment. Exercise patterns can also influence exposure.

Ambient concentration data tell only part of the story. Part of an individual's exposure includes contributions from non-ambient sources as well. These exposures include pollutants generated from indoor sources such as stoves. A significant body of recent research indicates that ambient PM concentrations provide a reasonable measure of the indoor concentrations of $PM_{2.5}$ to which an individual is exposed. Specifically, the EPA Criteria Document (1996) reports that "…in the absence of major non-ambient sources, total PM exposures to individuals tracked through time were highly correlated with ambient PM concentrations". In general, analysis of ambient data indicates that the gradient of fine particle concentrations is flat over urban areas. Therefore, it appears reasonable to assign monitored $PM_{2.5}$ concentration values as an individual's exposure index.

MEASUREMENT ERROR

The interpretation of the statistical association between ambient PM concentrations and mortality rates is hampered by possible measurement error in the pollution variable. One source of measurement error was mentioned above. That is the use of ambient concentrations to measure personal exposure. There are also other potential sources of measurement error that can affect the statistical results of an epidemiological analysis. Examples include:

- The use of limited data. Ambient data may be available for only selected periods of time. When data are taken from ambient monitors that are maintained for regulatory purposes,

[9] There is some research which links measured PM back to the originating source. We do not have such data for the set of studies included in the current analysis.

only minimum criteria concerning data availability are required. Data may be collected on a less than daily basis and monitors may not report data due to operational or maintenance downtime. Samples of observed (monitored) data used to form statistics might therefore be biased.

- The statistical measure used. A variety of statistics can represent the distribution of $PM_{2.5}$ concentrations over a specified period. For example, if a year is the temporal unit for air pollution data, one could summarize the observed (monitored) data within the year by the calculation of a mean value, by the median value or by some percentile observation of the underlying distribution. The latter measure index is most often used to summarize gaseous pollutant concentrations. Mean or median values are most often used for PM. Existing epidemiological studies have used both measures. While neither measure alone is inappropriate, analysts should maintain consistency between the measure used in an existing epidemiological study and the data used to evaluate the epidemiological function for analysis purposes.

- The time units of the pollution variable(s). The unit of time used to characterize air pollution concentrations affects the interpretation of the type of mortality outcome being evaluated. For example, a 24-hour concentration measure that is associated with daily mortality explains associations between acute exposure and acute mortality. Longer-term measures (e.g., annual values) are best thought of as providing insights into more chronic effects of exposure. There remains the question of how best to describe the temporal exposure of an individual in chronic exposure situations. Is a single year average value appropriate? Should there be separate pollution variables for a multi-year period? Is a lag structure appropriate?

- The spatial units of the pollution variable(s). Is it reasonable to use a statistic from a single ambient monitor to represent exposure for the entire population of a city? If data exists from multiple monitors within a city, how might that data be combined to provide a scalar measure of exposure? Is a city-wide average the proper spatial dimension?

INDIVIDUAL VERSUS ECOLOGIC DATA

The prospective cohort studies emphasized in this analysis have a design feature that sets them apart from other study frameworks. The observations include data that are specific to members of the cohort. Vital status, educational attainment. cigarette consumption, occupational exposure, age, ethnicity and other factors are specific to a cohort member. These data differ from the data used in many previous studies. Those studies often relied on community level data, also referred to as ecologic data, to define the statistical observations. There is one exception, the exposure variable. It would be too costly to collect individual data on exposure with personal dosimeters for sample sizes common to the cohort studies. As a result, the average concentrations for an area are assumed to apply to all individuals residing in the area. Typically, no adjustments are made to differentiate individual characteristics or behavioral actions. Apart

from measurement error issues, a concern has also been raised on the efficacy of using variables of different spatial dimension in the statistical analysis.

EXTRAPOLATION OF RESULTS

There is a range of PM concentrations implied by the geographic areas included in an underlying statistical analysis. This range defines a valid range for extrapolation in an economic benefit analysis. Extrapolation outside the range of observational data may introduce error since the estimated risk relationship is not based on the extrapolated values. In the SAMI application, there is another problem. The modeled concentration levels appear to understate monitored values. As a result, there may be the appearance of invalid extrapolation.

SUMMARY OF SENSITIVITY ISSUES

It seems appropriate to end this section on "controversies" in applied risk analysis with words used in the introductory paragraphs of the section. Estimation of an unbiased measure of risk is a difficult task. The art of risk assessment as applied to environmental problems has matured over time. But there is still room for growth. As a result, opinions can vary about the estimation and interpretation of mortality risk. Individuals must weigh the value of results obtained in "conventional" analyses against remaining uncertainties associated with unresolved (unresolvable?) questions. There is likely no single correct conclusion.

This section and earlier sections should be considered together to assist the reader in forming a personal judgement about quantitative risk analysis. In addition, a body of science exists outside the framework of epidemiological analysis. This other evidence, while not covered here, provides another basis for decisions.

STATISTICAL LIVES SAVED FOR CORE MODELS

CALCULATION METHOD

The procedure used to estimate Statistical Lives Saved is written as:

$$LS_{t,c,s} = \left(\frac{Mort_c}{1000}\right) \bullet Pop_{1999,c} \bullet E\left[\frac{Pop_{t,c}}{Pop_{1999,c}}\right] \bullet \left(1 - e^{-\frac{\ln(RR)}{Range} \bullet \Delta PM2.5_{t,c,s}}\right),$$

Where: $LS_{t,c,s}$ is Statistical Lives Saved at time t, in county c, for strategy s.
$\Delta PM2.5_{t,c,s}$ is the county average change in $PM_{2.5}$ at time *t* for county *c,* for strategy *s,*
RR is relative risk based on the selected model,
Range is the range in $PM_{2.5}$ in the selected risk study,
$Mort_c$ is the number of deaths from all causes per 1,000 in county *c* in 1994, and
$Pop_{t,c}$ is population for county *c,* in year *t*.
$E\left[\frac{Pop_{t,c}}{Pop_{1999,c}}\right]$ is the expected growth rate in county *c* population between 1999 and year *t.*

INPUT DATA

Key data elements of the quantitative analysis are:

- Annual average $PM_{2.5}$ concentrations for 12x12 and 24x24 grid cells in the modeled SAMI region. The SAMI Atmospheric Modeling group provided the data via spreadsheet. Location information (i.e., latitude and longitude) and a reference map were also available. The data represent predicted annual average values at the center point of each grid cell. Annual averages were computed using a weighting methodology proposed by the Atmospheric Modeling team. The methodology used weights computed for two Class I areas to convert daily modeled values (for 22 days) to annual averages. The weights for Shenandoah National Park were used for cells north of the North Carolina/Virginia border. The weights for Great Smokey National Park were applied elsewhere.

A series of steps was then taken to further develop the data for use in our analysis. Specifically:

- We eliminated grid cells that fell outside the SAMI region (e.g., Pennsylvania)
- We eliminated SAMI counties along the Atlantic Coast. This was done in recognition of the possible difference in modeling outcomes at coastal locations.
- We assigned grid cells to counties based on a minimum distance criterion. That is, a cell was assigned to a county such that the distance between the cell center point and the county center point was minimized.
- We reviewed the cell data to determine areas of significant deviation in values within a county area.

- We computed a size-weighted county average value for $PM_{2.5}$ with the weights reflecting the two different sizes of the grid cells.
- We noted that coverage of the 8 State SAMI region was about 85%. Some counties did not appear in the database because they were outside the 12x12 or 24x24 grid cell areas. These excluded areas are: southern Alabama and Georgia and western Tennessee and Kentucky. Figure 3, below, outlines the study area, with included areas in shaded colors.

The availability of modeled data at a detailed geographic resolution adds significant credibility to the analysis. However, it should be noted that the atmospheric model developed for SAMI was designed primarily for predicting strategy impacts in the Class I areas. In addition, the model was not intended to report annual average concentrations. The method used by the modelers to generate the data needed for this analysis is subject to uncertainty.

- County population data obtained from: (CO-99-1) County Population Estimates for July 1, 1999 and Population Change for July 1, 1998 to July 1, 1999. Source: Population Estimates Program, Population Division, U.S. Census Bureau, Washington, DC 20233
  Contact: Statistical Information Staff, Population Division, U.S. Census Bureau (301-457-2422)

- County mortality data obtained from: CDC National Center for Health Statistics as reported in the University of Oregon's County Data Center database. Data reflect mortality per 1,000 population by county in 1994. This baseline mortality rate is assumed to remain constant over time.

- Growth rate projections for population are taken from files provided by SAMI's emission inventory contractor.

- Relative risk values and associated changes in $PM_{2.5}$ are the core model results reported earlier.

- Results are summarized as "Statistical Lives Saved". Calculations are made for changes in exposure between OTW (also referred to as A2) (baseline) and two strategies (BB (B3) and BWC(B1)). Results are presented for 2010 and 2040 for the SAMI region and by State.

To illustrate the procedure used to compute the OTW versus BWC estimates in 2040, consider Autauga County, Alabama. The values of these factors for Autauga County are:

$LS_{2040,c,s}$	8.8
$\Delta PM2.5_{t,c,s}$	14.096−12.277 = 1.819
RR	1.12 (temporary assumed value)
Range	24.5 (temporary assumed value)
$Mort_c$	9
$Pop_{1999,c}$	43,140

$E\left[\frac{Pop_{2040,c}}{Pop_{1999,c}}\right]$	1.35995

$$LS_{2040,Autauga,s} = \frac{9}{1000} \times 43140 \times 1.35995 \times \left(1 - e^{-1.585 \times \frac{\ln(1.12)}{24.5}}\right)$$

$$LS_{2040,Autauga,s} = 528.014 \times \left(1 - e^{-0.0073317}\right) = 528.014 \times 0.007108 = 3.857$$

SUMMARY OF CORE RESULTS – STATISTICAL LIVES SAVED

Table 8: Krewski/Pope: Mean Measure of $PM_{2.5}$, RR=1.12; Range of $PM_{2.5}$ in study = 24.5
Point Estimates of Lives Saved for individuals ages 30 and over, by State, Strategy and Year

STATE	2010 (BWC-OTW)	2010 (BB-OTW)	2040 (BWC-OTW)	2040 (BB-OTW)
Alabama	113	387	305	537
Georgia	187	692	536	1029
Kentucky	63	303	94	307
North Carolina	247	911	624	1119
South Carolina	85	360	201	430
Tennessee	169	578	426	754
Virginia	108	381	331	561
West Virginia	31	123	80	168
Total	1002	3734	2598	4905

Table 9: Krewski/Pope: Mean Measure of $PM_{2.5}$, RR=1.06; Range of $PM_{2.5}$ in study = 24.5
Low Estimates of Lives Saved for individuals ages 30 and over, by State, Strategy and Year

STATE	2010 (BWC-OTW)	2010 (BB-OTW)	2040 (BWC-OTW)	2040 (BB-OTW)
Alabama	58	200	158	278
Georgia	96	358	277	533
Kentucky	33	156	49	158
North Carolina	127	471	322	579
South Carolina	44	186	103	222
Tennessee	87	299	220	390
Virginia	55	196	171	290
West Virginia	16	63	41	86
Total	516	1930	1340	2536

Table 10: Krewski/Pope: Mean Measure of $PM_{2.5}$, RR=1.19; Range of $PM_{2.5}$ in study = 24.5
High Estimates of Lives Saved for individuals ages 30 and over, by State, Strategy and Year

STATE	2010 (BWC-OTW)	2010 (BB-OTW)	2040 (BWC-OTW)	2040 (BB-OTW)
Alabama	174	591	467	819
Georgia	286	1054	819	1567
Kentucky	97	464	145	469
North Carolina	378	1389	955	1705
South Carolina	130	549	307	656
Tennessee	259	881	651	1150
Virginia	165	582	507	858
West Virginia	47	188	123	257
Total	1536	5698	3973	7481

Table 11: Krewski/Dockery: Mean Measure of $PM_{2.5}$, RR=1.28; Range of $PM_{2.5}$ in study = 18.6
Point Estimates of Lives Saved for individuals ages 25 and over, by State, Strategy and Year

STATE	2010 (BWC-OTW)	2010 (BB-OTW)	2040 (BWC-OTW)	2040 (BB-OTW)
Alabama	361	1,215	964	1,680
Georgia	608	2,212	1,728	3,278
Kentucky	203	966	302	977
North Carolina	794	2,883	1,995	3,540
South Carolina	270	1,129	637	1,351
Tennessee	541	1,823	1,355	2,376
Virginia	347	1,216	1,063	1,788
West Virginia	96	383	251	523
Total	3,220	11,826	8,295	15,513

Table 12: Krewski/Dockery: Mean Measure of $PM_{2.5}$, RR=1.09; Range of $PM_{2.5}$ in study = 18.6
Low Estimates of Lives Saved for individuals ages 25 and over, by State, Strategy and Year

STATE	2010 (BWC-OTW)	2010 (BB-OTW)	2040 (BWC-OTW)	2040 (BB-OTW)
Alabama	127	432	341	600
Georgia	214	791	612	1,177
Kentucky	71	341	106	345
North Carolina	279	1,030	706	1,265
South Carolina	95	403	225	482
Tennessee	190	650	479	849
Virginia	122	430	374	634
West Virginia	34	135	88	184
Total	1,130	4,212	2,932	5,535

Table 13: Krewski/Dockery: Mean Measure of $PM_{2.5}$, RR=1.49; Range of $PM_{2.5}$ in study = 18.6
High Estimates of Lives Saved for individuals ages 25 and over, by State, Strategy and Year

STATE	2010 (BWC-OTW)	2010 (BB-OTW)	2040 (BWC-OTW)	2040 (BB-OTW)
Alabama	579	1,927	1,537	2,657
Georgia	976	3,496	2,751	5,158
Kentucky	327	1,544	484	1,562
North Carolina	1,275	4,556	3,184	5,596
South Carolina	434	1,787	1,019	2,140
Tennessee	868	2,886	2,161	3,757
Virginia	558	1,939	1,702	2,847
West Virginia	155	614	404	838
Total	5,174	18,748	13,242	24,555

Table 14: Lipfert: Mean Measure of $PM_{2.5}$, Implied RR=0.9119; Implied Range of $PM_{2.5}$ in study = 15.09; Mean Estimates of Lives Saved for male veterans ages 25 and over with hypertension, by State, Strategy and Year

STATE	2010 (BWC-OTW)	2010 (BB-OTW)	2040 (BWC-OTW)	2040 (BB-OTW)
Alabama	-7	-23	-18	-32
Georgia	-10	-39	-30	-58
Kentucky	-3	-16	-5	-16
North Carolina	-13	-49	-33	-61
South Carolina	-5	-21	-12	-25
Tennessee	-9	-31	-23	-40
Virginia	-6	-21	-18	-32
West Virginia	-2	-7	-5	-10
Total	-54	-207	-143	-273

Tables 8 to 14 summarize the results of the analysis for each model and strategy. Estimates are reported by State to highlight geographical differences. For the two Krewski re-analyses, the models are represented by three tables. These tables show results for the point, lower 95% confidence value and upper 95% confidence value of the relative risk measure. Only the point estimate value is available for the Lipfert study.

It is appropriate to offer a comment about the negative results reported in Table 14 for the Lipfert study. One interpretation of these results is that higher concentrations of fine particles contribute to an increase in health status. But the lead author of the study provides cautionary notes about such an interpretation.

" There is also the question of the significant negative responses to air pollution. We note that they occur in each column of Table 9 and throughout Tables 6 and 7; thus, they are unlikely to be purely chance observations or due to one particular aspect of the model. However, these negative responses should not be interpreted as 'beneficial' effects of air pollution. The only conclusory statements that may be made are along the lines of: 'The members of this cohort who resided in counties with higher values of pollutant X tended to survive longer than those residing where concentrations were lower'." (Lipfert et al.,2000,p.65)

and

"This caveat also affects any causal attributions that might be hypothesized from the positive relationships. Thus, it is impossible to judge whether the final model used here is adequate or whether the use of additional mortality predictors, such as variables for diet and/or exercise (fn omitted) might have affected the pollution results even more." (Lipfert et al.,2000,p.65)

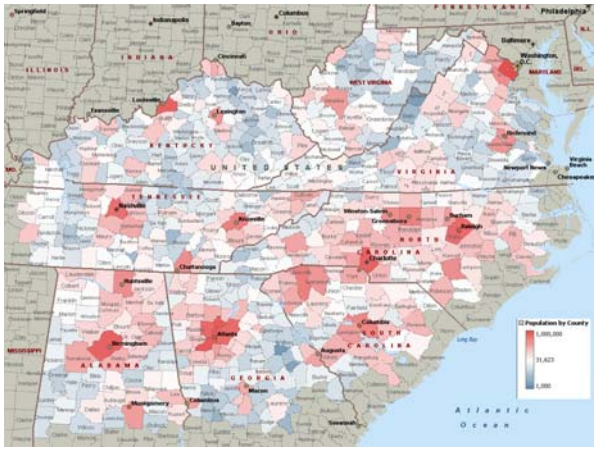
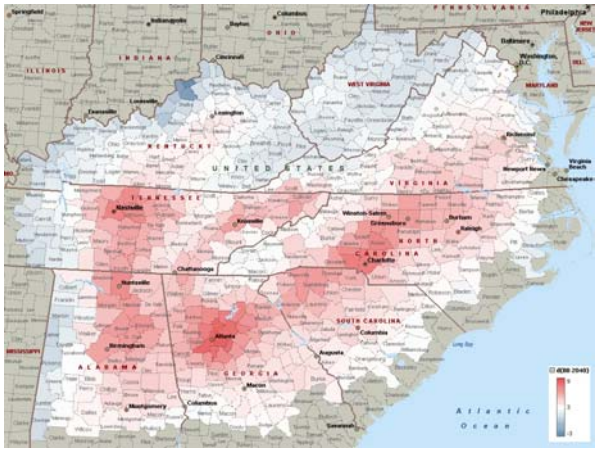


Exhibit 1: 1998 Population by county (log scale)
(Study area limited to shaded counties)

Figure 3 -- Population of SAMI Region

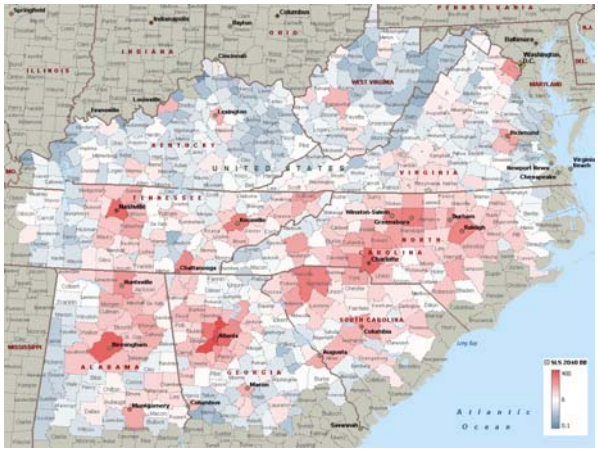


Estimated Change in PM2.5 in 2040 for Beyond Bold Relative to On-the-Way (Study area limited to shaded counties)

Figure 4 -- Change in Air Quality

Figures 3 to 5 are included to provide a more visual representation of the risk assessment results. Figure 3 shows the distribution of county population in the SAMI area in 1998. Note, that the shaded counties in the Figure are those in the analysis. Figure 4 provides a mapping of the estimated change in PM concentrations in 2040 for the A2 to B3 strategy. This Figure shows where the modeled air quality improvements occur. Figure 5 shows the distribution of statistical

lives saved across the SAMI region for one of the core models and the A2 to B3 strategy. There is a consistent pattern observed in the Figures. The air quality improvements are most pronounced in urban areas of the region and this result, in combination with higher population densities in the urban areas, contributes to the predicted statistical lives saved in urban areas.



SAMI Region – Estimated 2040 "Statistical Lives Saved" Based on Krewski/Pope for Beyond Bold Strategy Relative to On-the-Way. (log scale). Study area is shaded.

Figure 5 -- Estimated Statistical Lives Saved

Economic Valuation and Benefit Analysis

One of the purposes of this study is to estimate the economic value of benefits associated with different air quality regulatory strategies. We have done this by taking estimates of changes in $PM_{2.5}$ exposure associated with three different regulatory strategies, interacting these with estimates of the effect of $PM_{2.5}$ exposure on the risk of mortality from three different studies and forming estimates of the effects of these strategies on aggregate mortality risk. The reduction in mortality risk calculated in this way is usually measured in terms of "statistical lives saved." This section of our report discusses ways we can estimate the economic value of these "statistical lives saved."

Before we turn to the economic valuation, however, we believe it important to consider the meaning of the phrase "statistical lives saved." This phrase does not refer to saving an individual's life. It refers instead to reducing the risk of dying from a specific cause. This reduced risk is aggregated across many individuals. A reduction in risk of 1 in X (e.g., 1 in

10,000) when aggregated across X (10,000) persons is called a "statistical life saved." This is not the same as placing an economic value on a person's life. Even if all of the analysis and estimation is done perfectly, the value of a "statistical life saved" (VSL) is not equal to the value of a life – it is equal to X times the value of reducing risk by one chance in X.

Mortality Risk Valuation

In his comprehensive review of the literature, "The Value of Risks to Life and Health," Viscusi (1993) discusses three types of economic studies that attempt to place an economic value on a statistical life. The vast majority (23) of these studies are based on labor market studies relating hourly earnings to mortality and health risks.[10] He also reviews seven papers which estimate VSL by looking at risk – value trade-offs outside the labor market, and six papers based on survey methodology.

Since the publication of Viscusi (1993) there have been many studies that include estimates of VSL. The estimates of VSL found in these studies have not differed greatly from those cited by Viscusi. In early November, 2001, EPA sponsored a workshop titled "Economic Valuation of Mortality Risk Reduction: Assessing the State of the Art for Policy" in Silver Spring, Maryland. While the papers presented at this conference are not yet published, the results presented seemed consistent with those in Viscusi (1993). In fact, one of the presenters, Professor V. Kerry Smith, of North Carolina State University, made a particular note that his estimates ($6.5M) were in line with those found by Viscusi.

VSL Approach – Estimation Methods

The majority of studies cited in Viscusi (1993) are based on labor market data and follow the general method developed by Thayer and Rosen (1976). The basic method[11] is to estimate a wage equation of the form:

$$w_i = \alpha + \sum_{m=1}^{M} \Psi_m x_{im} + \gamma_0 p_i + \gamma_1 q_i + \gamma_2 q_i WC_i + u_i \quad ,$$

where: w_i is the annual wage of the i^{th} individual, the x_{im} are the personal and job related characteristics of the individual, p_i is the mortality risk associated with the person's job, q_i is the

[10] Interestingly one of these papers, Portney (1981), estimates VSL from data on air pollution and property values in Allegheny County, Pennsylvania. Viscusi reports Portney estimates of VSL for a 42 year old male as $800,000 in 1990 dollars. Portney's estimates are substantially smaller for older cohorts.

[11] The notation used here is taken directly from Viscusi (1993).

non-fatal job risks of the job, WC_i is the workers' compensation benefits payable if the worker is unable to work, and the u_i is a random error term. The Greek letters represent the parameters of the relationship that are to be estimated from the data. The estimated coefficient of mortality risk, γ_o, provides the way to estimate VSL. In particular the coefficient of mortality risk measures the average additional annual wage associated with a one unit change in mortality risk. This is the estimated value for VSL.

To illustrate assume that γ_o was estimated as \$6M. This implies that a decrease in mortality risk of 1 in 10,000 is equal to \$6,000,000/10,000, or \$600 per year in annual earnings. If 10,000 persons were exposed to a risk reduction of 1 in 10,000, we would expect one less death. If 10,000 persons receive the reduction in risk, then the benefits received would be \$600 time 10,000 or \$6,000,000. Thus the value of a statistical life would be estimated to be \$6M.

This illustration is designed to make it clear that VSL does not refer to the value of a particular individual's life, but rather to the aggregate value that many persons would be willing to pay for small mortality risk reductions.

VSL Approach – Some Concerns

The estimates of VSL cited in Viscusi (1993) are based on many different sets of data. Some of the studies are based on national samples of workers from across the spectrum of industries and occupations and others from more narrow samples.[12] The estimates derived from each of these studies are most appropriately applied to the populations the samples are taken from. For example, to apply VSL estimates based on a sample of young men to the general population requires us to believe that there are no significant differences between the way young men value mortality risk and the way the general population values mortality risk, or to believe at least that the model had been designed in such a way that these differences were adequately accounted for in the x_{im} variables. The risks involved (e.g., voluntary vs. involuntary) and the nature of the risk (e.g., latency, disutility) are also considerations.

Since VSL is an average measure specific to the population used to estimate the risk/value equation, we need to consider how the characteristics of the population in the SAMI region may differ from those populations the VSL estimates pertain to. In particular we need to

[12] For example, Arnould and Nichols (1983) uses a national sample from the US Census, while Brown (1980) uses a sample of young men from the National Longitudinal Survey of Young Men.

consider possible differences in age, health, race and gender characteristics, because there is ample evidence to indicate that these factors influence mortality risk.

In addition to the uncertainty associated with the population characteristics of the studies and the populations in the SAMI region, there are concerns that traditional VSL estimates might have un upward bias. Shogren and Stamland (2001) argue that wage-risk studies contain an upward bias because these studies cannot account for differences in the skill levels that lead workers to self-select into hazardous occupations. EPA's Science Advisory Board (1999) shares those concerns.

Several Government Agencies use VSL for Cost/Benefit Analysis.

Many governmental agencies use VSL as part of their regulatory process. There seems to be little agreement between these agencies as to what VSL estimates to use. In fact, the Food Safety and Inspection Service (FSIS) and the The Economic Research Service (ERS) of the Department of Agriculture, The Food and Drug Administration's Center for Food Safety and Applied Nutrition (CFSAN) and the Centers for Disease Control and Prevention (CDC) of the Department of Health and Human Services, and The Environmental Protection Agency (EPA) held a joint conference on the topic in September, 2000. The conference proceedings state:

> The different agencies use many different methods to estimate the reductions in risk associated with their policies; they also use different methods to assign dollar values to the human health benefits from the reduced risk. Because each agency uses a different method, it is difficult to compare programs across agencies. If there were a consensus approach to assigning values to risk reductions and each agency used it, risk managers could compare the effects of different agencies' programs.
>
> …
>
> As a first step toward generating a common approach, several agencies sponsored a conference on valuing the health benefits of food safety. The conference was a first step toward forging a consensus across agencies on how to value improvements in food safety. Most of the issues discussed applied to general health and safety issues in addition to food safety.

Table 15 summarizes some of the disparate VSL values used by three government agencies.

Table 15: VSL Estimates used by Three Government Agencies

Agency	VSL Value	Standard Error of VSL	Year of Measure
EPA	$6.3M	$1.6M[13]	1999
USDA	$15k to $2.0M (Depending on Age)	NA	1996
DOT	$3.0M	NA	1996

Source: Kenkel (2000), pp 2-9.

The FDA also computes a per life year saved version of VSL referred to as a quality-adjusted life year or QALY. According to Kenkel:

> The quality adjusted life year, or QALY, has emerged as a standard measure of effectiveness, and the cost per QALY saved has been estimated for a wide range of health interventions (Gold et al., 1996). The QALY approach not only incorporates the quantity of life or years of life extension from an intervention, but also the quality of life, based on individuals' preferences over different health states. For example, a year of life with a serious illness might be weighted as being as valuable as 0.7 of a year of life with perfect health. Placing a monetary value on a QALY allows the health effects to be monetized, converting any cost-effectiveness analysis into a cost-benefit analysis.

While the QALY methods are of interest, we do not believe they are sufficiently well developed and accepted in the literature to be used in this study.

Recommendations for VSL

There are many differences in methods used by researchers estimating the economic value to be assigned small reductions in risk shared by large populations. Government agencies regularly use these estimates to compute the benefits accruing from their regulatory activities. Neither academics nor government agencies agree on how these benefits should be estimated. It is clear that there are benefits from reduced risk of mortality. It seems from the literature that, in 2000 dollars, these benefits are on the order of one to ten million dollars. We agree with several others that EPA's estimate of 6.3 million dollars (1999) is likely to be biased upward. We recommend that the value of additional $PM_{2.5}$ regulations imposed on the SAMI region be estimated with at least three alternative VSL values. Given the existing knowledge we recommend these values be one, three and six million dollars (year 2000 $).

We fully recognize that these are arbitrarily chosen values, but believe that they are reasonable values given the current state of knowledge.

[13] Implied value derived from cited range of VSL estimates.

Combining VSL with Statistical Lives Saved Estimates

The process we followed to place economic values on the statistical lives saved estimates are summarized by the following steps:

1. Choose year 2000 VSL as basis of calculations,
2. Adjust VSL for each year through 2040 by applying an estimate of the income elasticity for the willingness to pay to avoid premature death to forecasts of GDP increases,
3. Use Statistical Lives Saved (SLS) values for 2010 and 2040 to generate 2011 through 2039 SLS values assuming exponential growth,
4. Calculate the implied value of the statistical lives saved for 2010 through 2040,
5. Calculate the year 2010 present value (PV) of the dollar stream generated in step 4 based on a real discount rate of 5 percent. This calculation assumes that the statistical lives saved are uniformly distributed across each year of study,
6. Calculate the annualized value of the PV computed in step 5,
7. Discount the value in step 5 back to year 2000.

To illustrate the process, we take the statistical lives saved derived for the B1 strategy for the Krewski/Pope case:

VSL_{2000} = $1,000,000[14]
Income Elasticity = $\eta = 0.4$
SLS_{2010} = 1,002
SLS_{2040} = 2,598

The calculations are shown in Table 16.

[14] We use a value of $1,000,000 for VSL since this can easily be scaled.

Table 16 – Calculations of Annualized Value of SLS

Year	GDP Estimate	VSL_{year}	SLS_{year}	$VSL_{year} * SLS_{year}$	PV_{year}
2000	27.7	1,000,000			
2001	28.1	1,005,776			
2002	28.5	1,011,503			
2003	29	1,018,601			
2004	29.4	1,024,221			
2005	29.9	1,031,189			
2006	30.3	1,036,707			
2007	30.8	1,043,550			
2008	31.2	1,048,971			
2009	31.7	1,055,695			
2010	32.2	1,062,355	1,002.00	1,064,480,109	1,038,723,185
2011	36.6	1,120,422	1,034.33	1,158,888,735	1,076,997,555
2012	37.2	1,127,769	1,067.71	1,204,128,008	1,065,752,439
2013	37.9	1,136,258	1,102.16	1,252,338,354	1,055,640,538
2014	38.7	1,145,851	1,137.72	1,303,663,586	1,046,575,620
2015	39.5	1,155,326	1,174.44	1,356,857,426	1,037,408,962
2016	40.3	1,164,686	1,212.33	1,411,987,190	1,028,151,793
2017	41	1,172,778	1,251.45	1,467,675,802	1,017,811,411
2018	41.7	1,180,787	1,291.83	1,525,381,050	1,007,456,329
2019	42.3	1,187,583	1,333.52	1,583,664,198	996,142,988
2020	43.6	1,202,182	1,376.55	1,654,861,968	991,359,242
2021	43.6	1,202,182	1,420.97	1,708,260,725	974,617,357
2022	44.2	1,208,800	1,466.82	1,773,089,235	963,432,471
2023	44.8	1,215,363	1,514.15	1,840,241,222	952,305,170
2024	45.4	1,221,874	1,563.01	1,909,798,305	941,238,302
2025	46.01	1,228,420	1,613.44	1,981,984,493	930,300,044
2026	46.62	1,235,001	1,665.50	2,056,899,161	919,488,900
2027	47.25	1,241,617	1,719.25	2,134,645,438	908,803,395
2028	47.88	1,248,268	1,774.72	2,215,330,355	898,242,067
2029	48.52	1,254,955	1,831.99	2,299,064,984	887,803,473
2030	49.17	1,261,678	1,891.10	2,385,964,599	877,486,188
2031	49.83	1,268,437	1,952.13	2,476,148,829	867,288,802
2032	50.50	1,275,233	2,015.12	2,569,741,825	857,209,921
2033	51.17	1,282,064	2,080.14	2,666,872,430	847,248,168
2034	51.86	1,288,932	2,147.26	2,767,674,360	837,402,181
2035	52.55	1,295,837	2,216.55	2,872,286,382	827,670,617
2036	53.26	1,302,779	2,288.07	2,980,852,508	818,052,144
2037	53.97	1,309,759	2,361.90	3,093,522,196	808,545,448
2038	54.69	1,316,775	2,438.12	3,210,450,552	799,149,232
2039	55.43	1,323,829	2,516.79	3,331,798,542	789,862,210
2040	56.17	1,330,921	2,598.00	3,457,733,221	780,683,114

The GDP forecasts through 2024 are based on those used in EPA's Heavy Duty Deisel Report. They were extended to 2040 by applying the 2023 to 2024 growth rate of approximately 1.34 percent to 2025 through 2040.

The value of VSL_{2001} is computed as: $VSL_{2001} = VSL_{2000} + VSL_{2000} \cdot \left(\frac{GDP_{2001}}{GDP_{2000}} - 1 \right) \eta$. In words we scaled the previous year's *VSL* up by the income elasticity times the growth rate in *GDP*. For example, if *GDP* increased by 2 percent, we increased *VSL* by 0.8 percent.

To compute the SLS_{2011} we determined the growth rate for the period 2010 to 2040 and determined the annual compound growth rate. This growth rate was then applied to each year. The formula used was: $SLS_{year} = SLS_{year-1} \cdot \left(\frac{SLS_{2040}}{SLS_{2010}} \right)^{(year-2010)/30}$.

The next column is simply the product of the previous two columns. The last column calculates the discounted value of these statistical lives saved. There are three ways we could reasonably do this calculation: (1) we could assume all lives are saved at the beginning of each year, (2) we could assume all lives are saved at the end of each year, and (3) we could assume that the lives saved are uniformly distributed throughout the year. We have chosen the last and used a standard actuarial method to accomplish the discounting.[15] The formula used is: $\mathrm{PV}_{\mathrm{year}} = C_{year} \cdot \left(\frac{\ln(1+0.05)}{0.05} \right) \cdot (1+0.05)^{-(year-2010)}$, where *C* represents the value obtained in the previous column. The $\left(\frac{\ln(1+0.05)}{0.05} \right)$ term makes the adjustment to distribute lives saved over the year.

Summing the final column we obtain a value of $28,848,849,264 ($28.8 Billion). This is annualized to $1,805,370,612 ($1.8 Billion) by dividing by the factor: $\frac{1-(1+0.05)^{-31}}{\ln(1+0.05)}$. This value is then converted to $1,108,340,946 ($1.1 Billion) , year 2000 dollars, by multiplying it by $(1+0.05)^{-10}$. This is the method used to compute the value we report as the annualized value of statistical lives saved.

[15] See Bowers et al. (1995), pp xx.

In Table 17, we report core results for the three studies used throughout this report.

Table 17: Economic Valuation of Annualized Benefits per $1M (2000 $) VSL by Study and Strategy

Study Characteristics			Strategy	
Study	RR	Range	A2 to B1	A2 to B3
Krewski/Pope	1.06	24.5	$571,185,857	$1,581,344,432
	1.12		1,108,340,946	3,059,081,091
	1.19		1,697,091,178	4,667,080,450
Krewski/Dockery	1.09	18.6	1,250,351,340	3,451,222,219
	1.28		3,550,898,545	9,682,874,209
	1.49		5,688,222,810	15,340,654,963
Lipfert	0.9119	15.09	-60,329,233	-169,863,451

The values in this table represent annualized benefits per million dollars of VSL. To obtain estimates based on a VSL of three million dollars simply multiply the values in the last two columns by 3. Thus, the Krewski/Pope, A2 to B1, RR = 1.12, estimate would be $3,325,022,838 ($3.3 Billion).

Earlier in this report we commented on the interpretation of the estimates based on the Lipfert study. We agree with Lipfert's own assessment of the caution needed to use and interpret his mortality risk values. The following are taken from Lipfert (2000):

> P. 65 " There is also the question of the significant negative responses to air pollution. We note that they occur in each column of Table 9 and throughout Tables 6 and 7; thus, they are unlikely to be purely chance observations or due to one particular aspect of the model. However, these negative responses should not be interpreted as 'beneficial' effects of air pollution. The only conclusory statements that may be made are along the lines of: 'The members of this cohort who resided in counties with higher values of pollutant X tended to survive longer than those residing where concentrations were lower'."
>
> P65. "This caveat also affects any causal attributions that might be hypothesized from the positive relationships. Thus, it is impossible to judge whether the final model used here is adequate or whether the use of additional mortality predictors, such as variables for diet and/or exercise (footnote omitted) might have affected the pollution results even more."

P. 69 “Finally, special attention must be given to significant negative associations between pollution and mortality, which may be indicators of confounding or an incomplete model specification. Obviously, such problems could affect the positive associations as well. In a cross-sectional analysis, the gradients pertain first and foremost to the characteristics of places; extension of those findings to the air pollution exposures of one-time residents of those places may or may not be justified.”

Bibliography

Arnould, R and Grabowski H. "Auto Safety Regulations: An Analysis of Market Failure," *Bell Journal of Economics*, Spring 1981, 12(1), pp. 27-48.

Bowers, N et al. *Actuarial Mathematics, 2ed*. Society of Actuaries, Schumsburg Ill. 19xx.

Brown, C. "Equalizing Differences in the Labor Market," *Quarterly Journal of Economics,* Feb. 1980, 94(1), pp. 113-34.

Gold, M., D. Patrick, G. Torrance, et al. 1996. "Identifying and Valuing Outcomes." Chapter 4 in *Cost-Effectiveness in Health and Medicine*, M. Gold, J. Siegel, L. Russell, and M. Weinstein, editors. Oxford University Press: New York. Pp. 82-134.

Kenkel, D, "Using Estimates of the Value of Statistical Life in Evaluating Regulatory Effects," *Valuing the Health Benefits of Food Safety* conference paper , College Park Maryland, September 14, 2000.

Kuchler, F and E Golan. 1999. *Assigning Values to Life: Comparing Methods for Valuing Health Risks*. U.S. Department of Agriculture, Economic Research Service, AER-784.

Kuchler, F, *Valuing the Health Benefits of Food Safety*, USDA, Miscellaneous Publication Number 1570, April 2001.

Letter to C. Browner, Administrator, US EPA, "The Clean Air Act Amendments (CAAA) Section 812 Prospective Study of Costs and Benefits (1999): Advisory by the Advisory Council on Clean Air Compliance Analysis: Costs and Benefits of the CAAA," October 29, 1999.

Portney, P., "Housing Prices, Health Effects, and Valuing Reductions in the Risk of Death," *Journal of Environmental Economics and Management*, March 1981, 8(1), pp. 72-8.

Shogren, J. and T. Stamland. May, 2001. "Skill and the Value of Life," Working paper, University of Wyoming.

Thayler, R and Rosen, S. "The value of Saving a Life: Evidence from the Market," in *Household production and consumption.* Ed. Terleckyj. Cambridge: NBER, 1976, pp. 265-98.

Viscusi, W K, "The Value of Risks to Life and Health," *Journal of Economic Literature*. 31 (4) (1993): 1912-1946.