Proposed Methodology for Particulate Matter Risk Analyses for Selected Urban Areas

January 2002

Prepared for

Office of Air Quality Planning and Standards
U.S. Environmental Protection Agency
Research Triangle Park, NC

Prepared by

Ellen Post
Ben Apelberg
Leland Deck
Don McCubbin

Work funded through
Contract No. 68-D-98-001
Work Assignments 3-51 and 4-65

Harvey Richmond, Work Assignment Manager
Nancy Riley, Project Officer

DISCLAIMER

This report is being furnished to the U.S. Environmental Protection Agency by Abt Associates Inc. in partial fulfillment of Contract No. 68-D-98-001, Work Assignment Nos. 3-51 and 4-65. Some of the preliminary work for this report was completed under Work Assignments 1-36 and 2-46 of the same contract. The opinions, findings, and conclusions expressed are those of the authors and are not necessarily those of the Environmental Protection Agency. This draft report is being circulated for review by the Clean Air Scientific Advisory Committee and the general public. All comments and inquiries should be addressed to Mr. Harvey Richmond, U.S. EPA, Office of Air Quality Planning and Standards, C539-01, Research Triangle Park, North Carolina 27711.

Table of Contents

List of Exhibits

PROPOSED METHODOLOGY FOR PARTICULATE MATTER RISK ANALYSES FOR SELECTED URBAN AREAS

1. Introduction

As required by the Clean Air Act, the U.S. Environmental Protection Agency (EPA) periodically reviews the national ambient air quality standards (NAAQS) for particulate matter (PM). As a result of the last review of the PM NAAQS completed in 1997 (62 FR 38652, July 18, 1997), EPA added new standards for $PM_{2.5}$, referring to particles with a mean aerodynamic diameter less than or equal to 2.5 μm, in order to address concerns about the fine fraction of inhalable particles. The existing PM_{10} standards, referring to particles with a mean aerodynamic diameter less than or equal to 10 μm, were originally adopted in 1987; they were retained with minor revisions in 1997 for the purpose of regulating the coarse fraction of inhalable particles.[1] The new $PM_{2.5}$ standards included: an annual standard of 15 $\mu g/m^3$, based on the 3-year average of annual arithmetic mean $PM_{2.5}$ concentrations from single or multiple community-oriented monitors; and a 24-hour standard of 65 ug/m^3, based on the 3-year average of the 98^{th} percentile of 24-hour $PM_{2.5}$ concentrations at each monitor in an area. These standards were based primarily on a large body of epidemiological evidence relating ambient PM concentrations to various adverse health effects.

As part of its last review, EPA's Office of Air Quality Planning and Standards (OAQPS) sponsored risk analyses for two sample urban areas, Philadelphia County and Los Angeles County, to assess the risks associated with current PM levels and the effects of alternative PM standards on reducing estimated health risks attributable to PM (U.S. EPA, 1996b, pp. VI-1 - VI-60; Abt Associates Inc., 1996; and Abt Associates Inc., 1997a,b. See also Deck et al., 2001 and Post et al., 2001 for published articles describing the risk analysis methodology used in the 1996-1997 analyses). As discussed in the Federal Register notice explaining the Administrator's decision to set new $PM_{2.5}$ standards (62 FR 38656), EPA did not rely on the quantitative results of these risk analyses in setting the levels of the standards because it concluded that the significant uncertainties inherent in the analyses precluded such reliance. Rather, EPA used the analyses in a more limited qualitative manner.

The next periodic review of the PM NAAQS is now underway. EPA is currently completing the process of assessing the latest available PM health effects literature. The latest draft of this assessment is contained in the March 2001 second external review draft of the *Air Quality Criteria for Particulate Matter* (U.S. EPA, 2001a) *(*hereafter 2001 draft PM CD). A third external review draft of the *Air Quality Criteria for Particulate Matter* is expected to be

[1] In May 1999, in response to challenges filed by industry groups, the U.S. Court of Appeals for the District of Columbia Circuit vacated the revised PM_{10} standards on the basis that PM_{10} is an "arbitrary indicator for coarse particulate pollution" American Trucking Associations v. EPA, 175 F. 3d 1027, 1053-55 (D.C. Cir. 1999). The 1987 PM_{10} standards remain in effect.

released in early 2002 for review by the Clean Air Scientific Advisory Committee (CASAC) and general public. The 2001 draft PM CD includes an evaluation of the scientific evidence on the health effects of PM, including information on exposure, physiological mechanisms by which PM might damage human health, and an evaluation of the epidemiological evidence including reported concentration-response (C-R) relationships.

At the time of the last PM CD (U.S. EPA, 1996a), a number of health studies indicated differences in health effects between fine and coarse fraction particles, and suggested that serious health effects, such as premature mortality, were more closely associated with fine fraction particles. The new studies, summarized in Chapter 6 of the 2001 draft PM CD continue to show associations between serious health effects, including premature mortality, and ambient $PM_{2.5}$ concentrations. In the last PM NAAQS review, there were only a limited number of studies that assessed the relationship between ambient $PM_{2.5}$ and various health effects. As shown in Exhibits C.1, C.2, and C.4 in Appendix C, there are significantly more studies available today that address the relationship between ambient $PM_{2.5}$ levels and significant health effects, including increased mortality associated with short- and long-term exposures, increased hospital admissions, and increased respiratory symptoms. As discussed more fully in Sections 3 and 4, these new studies include single-city studies in a variety of locations across the United States and Canada, as well as some multi-city studies. The health effects studies summarized in Chapter 6 of the 2001 draft PM CD also offer new evidence indicating possible associations between coarse fraction PM and health effects.

OAQPS also has released a preliminary draft Staff Paper, *Review of the National Ambient Air Quality Standards for Particulate Matter: Policy Assessment of Scientific and Technical Information* (U.S. EPA, 2001b)(hereafter preliminary draft PM SP) which identifies the key policy-relevant scientific information contained in the 2001draft PM CD. When finalized, the OAQPS Staff Paper will evaluate the policy implications of the key studies and scientific conclusions presented in the Criteria Document as well as various analyses (e.g., air quality, risk) summarized in the Staff Paper. The final Staff Paper will present factors relevant to the evaluation of the current primary (health-based) NAAQS, as well as staff conclusions and recommendations of options for the Administrator to consider. Consistent with EPA's conclusion from the last review that fine and coarse fraction PM be considered separately, the final OAQPS Staff Paper is expected to include recommendations for both $PM_{2.5}$ and $PM_{10-2.5}$ standards.

Therefore, the proposed PM risk analyses will focus on assessing risk associated with $PM_{2.5}$ and, to the extent appropriate, $PM_{10-2.5}$.[2] The proposed PM risk analyses will be based on

[2]Coarse particle concentrations have been measured directly using a dichotomous sampler or by subtraction of particles measured by a $PM_{2.5}$ sampler from those measured by a co-located PM_{10} sampler. This measurement is an indicator for the fraction of coarse-mode thoracic particles (i.e., those capable of penetrating to the tracheo-bronchial and the gas-exchange regions of the lung).

the health effects evidence assessed in the next draft of the PM CD, including many of the studies assessed in the prior PM CD, *Air Quality Criteria for Particulate Matter* (U.S. EPA, 1996a), and considered in the previous risk analyses. The current recommendations concerning health effects and studies to include in the PM risk analyses are based on the current draft PM CD and are, therefore, subject to revision once the next draft is available.

Decisions on whether EPA will propose and carry out risk analyses for $PM_{10-2.5}$, and if so to what extent, have not yet been made. This document contains placeholder sections for the methodological approach that would be taken to carry out risk analyses for $PM_{10-2.5}$ that will be added, if appropriate, pending OAQPS and Abt Associates review of the evaluation of $PM_{10-2.5}$ health effects evidence contained in the next draft of the PM CD.

The goals of the proposed PM risk analyses are: (1) to develop a better understanding of the influence of various inputs and assumptions on the risk estimates and (2) to gain qualitative insights into the nature of the risks associated with exposures to PM. In addition, the planned risk analyses also will provide a rough sense of the potential magnitude of PM-related mortality and morbidity associated with current $PM_{2.5}$ levels and with attaining the current $PM_{2.5}$ NAAQS (as well as any potential alternative $PM_{2.5}$ standards identified as part of this review). Finally, if EPA judges it appropriate to proceed with risk analyses for $PM_{10-2.5}$, then these analyses would provide a rough sense of the potential magnitude of PM-related morbidity associated with current $PM_{10-2.5}$ levels and with attaining possible alternative $PM_{10-2.5}$ NAAQS.[3] EPA recognizes that the role of the risk analyses in this standards review will necessarily be limited by significant uncertainties, as discussed in Section 6 below, and does not plan to use the risk estimates as a basis for recommending selection among alternative standard levels.

Given the availability of additional urban locations with recent and sufficient $PM_{2.5}$ air quality data, and the publication of additional health effect studies in various locations in different regions of the United States, EPA proposes to expand somewhat the scope of its PM risk analyses to several additional urban areas, consistent with the goals of the assessment. Philadelphia and Los Angeles Counties, which were the only areas included in the prior risk analyses, would be included. As discussed in greater detail in Section 3, proposed additional areas to be included for short- and long-term exposure mortality for $PM_{2.5}$ risk analyses are as follows: Boston, Detroit, St. Louis, Phoenix, and San Jose. In addition, increased hospital admissions associated with $PM_{2.5}$ would be estimated for Detroit, Los Angeles, and Seattle and respiratory symptoms for Boston and St. Louis. Inclusion of these additional areas will allow EPA to explore potential geographic differences in risk estimates.

The basic question addressed in the first part of the risk analyses is of the following form:

[3]EPA does not plan to assess risks associated with PM_{10} for this review.

For a given human health endpoint (mortality, hospital admissions, etc.), what is the estimated annual incidence of the health endpoint that may be associated with "as is"[4] PM concentrations above background in these locations?[5]

The second part of the risk analyses estimates the risk reductions that would result if a specific set of PM standards were just met in the selected locations. The basic question addressed in this part of the analyses is of the following form:

For a given human health endpoint (mortality, hospital admissions, etc.), what is the estimated reduction in annual incidence in terms of percentage and absolute numbers associated with the reduction in PM concentrations that would be expected to result if a specified set of PM standards were just met?

The methods proposed to be used to estimate risks and risk reductions in the selected urban areas in the planned PM risk analyses are similar to the methods used in the previous PM risk analyses. An overview of these methods is presented in Section 2, including discussion of any significant differences in approach from the risk analyses conducted for the last review. Section 3 discusses the selection of proposed health endpoints and urban areas from a broader list of candidate health endpoints and locations to include in the risk analyses. Section 4 describes the proposed approach to selecting and using C-R functions from the broader candidate pool of C-R functions available. Section 5 presents baseline health effects incidence rates (i.e., the health effects incidence rates associated with "as is" PM levels) for each of the proposed assessment locations. Because the risk analyses must be carried out with incomplete information, it is necessary to make assumptions at several points in the analysis process. These assumptions and the various sources of uncertainty in the analyses are discussed in Section 6. Appendix A discusses the air quality data proposed to be used in the analyses. Appendix B describes analysis of historical air quality data relevant to the choice of air quality adjustment procedure for simulating attainment of alternative PM standards. Appendix C summarizes information related to the selection of proposed health effect endpoints, urban areas, and epidemiological studies from a broader pool of candidates for inclusion in the risk analyses.

---

[4]"As is" PM concentrations are defined here as a recent year of air quality.

[5]Consistent with the approach taken in the prior PM risk analyses, risks only will be estimated for concentrations exceeding "background" levels, where "background" is defined as the PM concentrations that would be observed in the U.S. in the absence of anthropogenic, or man-made, emissions of primary PM and precursor emissions of volatile organic compounds, nitrogen oxides, sulfur dioxide, and ammonia in North America. Thus, "background" for the purposes of the PM risk analyses includes PM from natural sources and transport of PM from sources outside of North America.

2. Overview of Methods

This section gives an overview of the methods proposed to be used in the risk analyses. Section 2.1 presents the basic structure of the risk analyses, distinguishing between its two parts (i.e., risk associated with "as is" PM levels (defined as a recent year of air quality) and risk reductions associated with just meeting the current or alternative PM standards) and identifying the basic information elements needed for the analyses. Section 2.2 addresses the approach for estimating risk associated with "as is" PM levels above background. Section 2.3 discusses issues involved in estimating risks associated with just meeting alternative PM standards. Estimation of background PM concentrations in the assessment locations is discussed in Section 2.4. The predominant functional form of the C-R functions used in the risk analyses, and the prediction of changes in health effects incidence associated with changes in ambient PM concentrations using these C-R functions is described in Section 2.5. Issues involved in the calculation of annual health effects incidence are discussed in Section 2.6. A brief discussion of issues surrounding baseline incidence rates is given in Section 2.7. An overview of the sources of uncertainty in the PM risk analyses and proposed ways to handle uncertainty, is discussed in Section 2.8. Finally, proposed sensitivity analyses are discussed in Section 2.9.

2.1. Basic structure of the risk analyses

The general approach used in both the prior and the planned PM risk analyses relies upon C-R functions which have been estimated in epidemiological studies. Since these studies estimate C-R functions using air quality data from fixed-site, population-oriented monitors, the appropriate application of these functions in PM risk analyses similarly requires the use of air quality data at fixed-site, population-oriented monitors. The general PM health risk model combines information about PM air quality for specific urban areas with C-R functions derived from epidemiological studies and baseline health incidence data for specific health endpoints and population estimates to derive estimates of the annual incidence of specified health effects attributable to PM concentrations. The analyses are conducted for both "as is" air quality and for air quality simulated to reflect attainment of alternative PM standards.

Each part of the proposed PM risk analysis can be characterized as estimating the change in the incidence of a given health effect resulting from a given change in PM concentrations. In the first part, the change is from "as is" PM levels down to background (or to the lowest measured level (LML) observed in the study, if it is higher than the estimated background level in the assessment location). In the second part of the risk analyses, the change is from "as is" PM concentrations to those concentrations that would result if a specified set of PM standards were just met in the assessment locations. The method used in both parts of the risk analyses is therefore basically the same. The important difference between the two parts is in the specified lower PM levels. In the first part, the lower PM level is background (or the LML in the study). In contrast, the lower PM levels in the second part of the risk analyses are the estimated PM

levels that would occur when a specified set of PM standards is just met in the assessment locations. The second part requires that a method be developed to simulate just meeting specified standards.

An overview of the major components of the proposed PM risk analyses discussed in this report is presented in Exhibit 2.1. The points in the risk analyses at which sensitivity analyses will be carried out are represented by diamonds. The planned sensitivity analyses (labeled in Exhibit 2.1 as s_k's) are described in Exhibit 2.5 below.

To estimate the change in the incidence of a given health effect resulting from a given change in ambient PM concentrations in an assessment location, the following analysis inputs are necessary:

- **Air quality information** including: (1) "as is" air quality data for $PM_{2.5}$ from population-oriented monitors in the assessment location, (2) estimates of background $PM_{2.5}$ concentrations appropriate to this location, and (3) a method for adjusting the "as is" data to reflect patterns of air quality change estimated to occur when the area just meets various alternative standards. To carry out a $PM_{10-2.5}$ risk analysis, "as is" data for $PM_{10-2.5}$ (currently obtained by subtracting $PM_{2.5}$ concentrations from PM_{10} data from co-located population-oriented monitors) and estimates of background $PM_{10-2.5}$ would be required.

- **Concentration-response function(s)** which provide an estimate of the relation between the health endpoint of interest and PM concentrations (preferably derived in the assessment location, although functions estimated in other locations can be used at the cost of increased uncertainty -- see Section 6.1.3). For $PM_{2.5}$ C-R functions are available from epidemiological studies for both short- and long-term exposures.

- **Baseline health effects incidence rate and population**. The baseline incidence rate provides an estimate of the incidence rate (number of cases of the health effect per year, usually per 10,000 or 100,000 general population) in the assessment location corresponding to "as is" PM levels in that location. To derive the total baseline incidence per year, this rate must be multiplied by the corresponding population number (e.g., if the baseline incidence rate is number of cases per year per 100,000 population, it must be multiplied by the number of 100,000s in the population).

The risk analysis procedure described in more detail below is diagramed in Exhibit 2.2 for analyses based on short-term exposure studies and in Exhibit 2.3 for analyses based on long-term exposure studies.

Exhibit 2.1 Major Components of Particulate Matter Health Risk Analyses

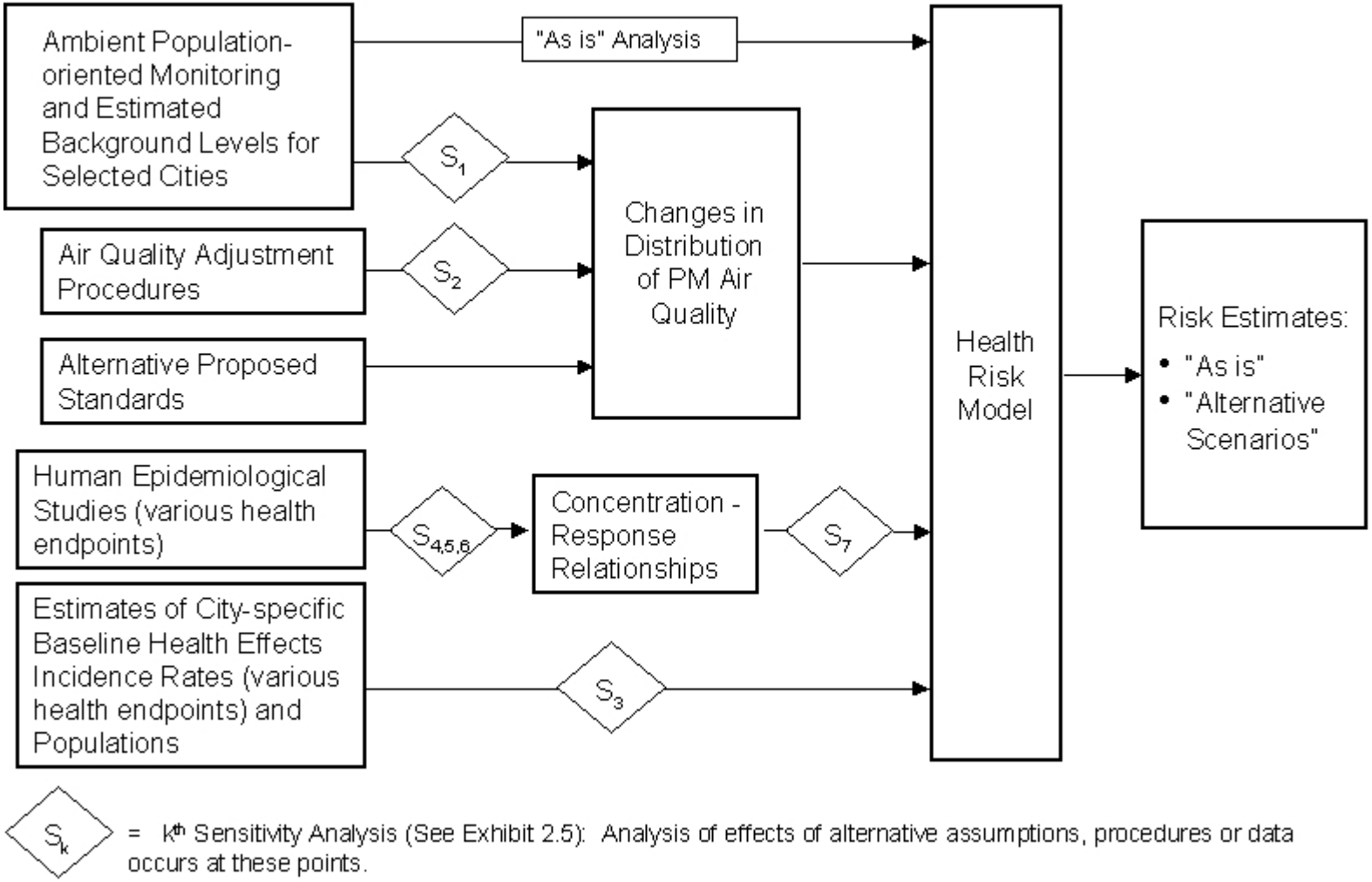


Exhibit 2.2 Flow Diagram of Proposed Risk Analyses for Short-Term Exposure Studies

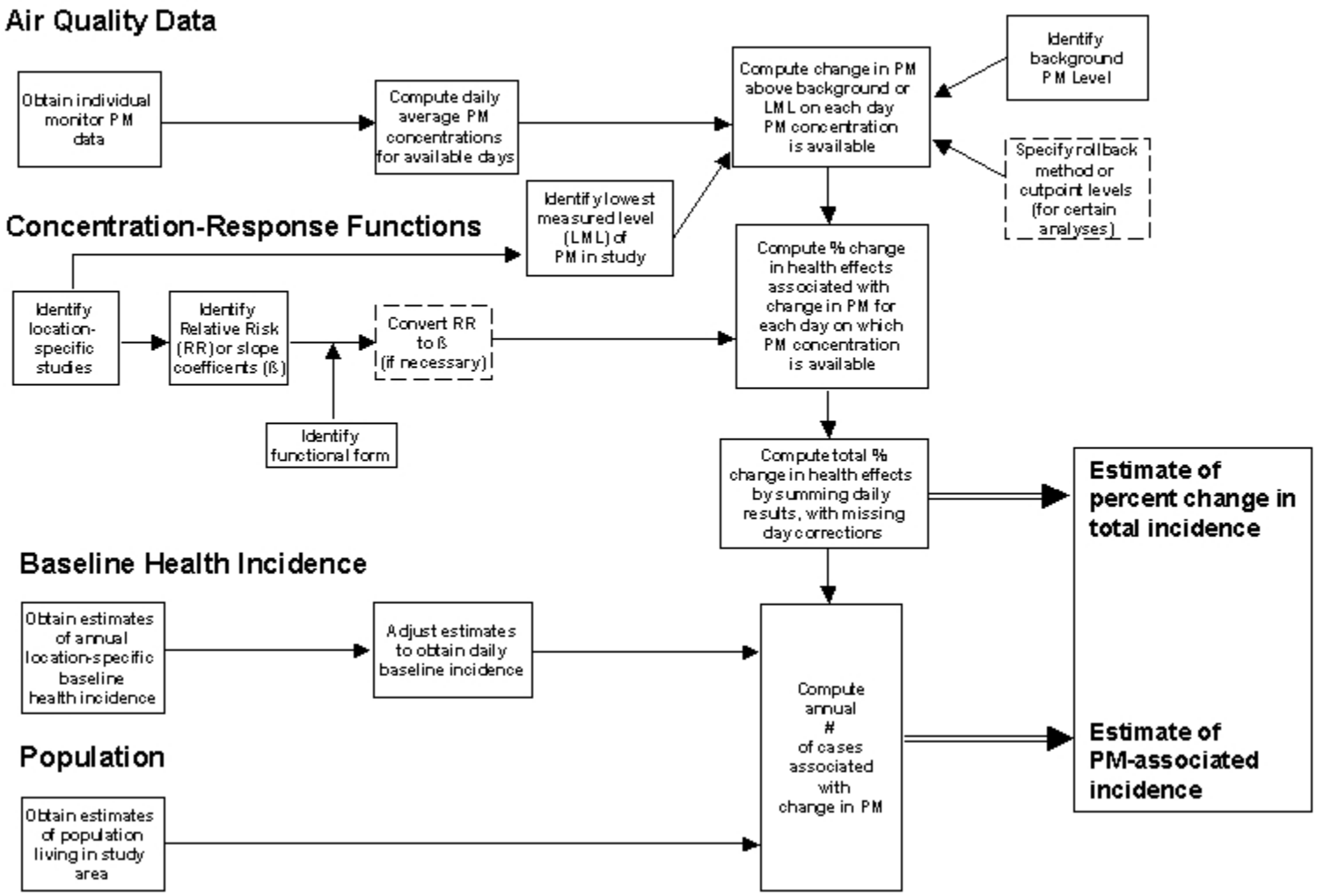
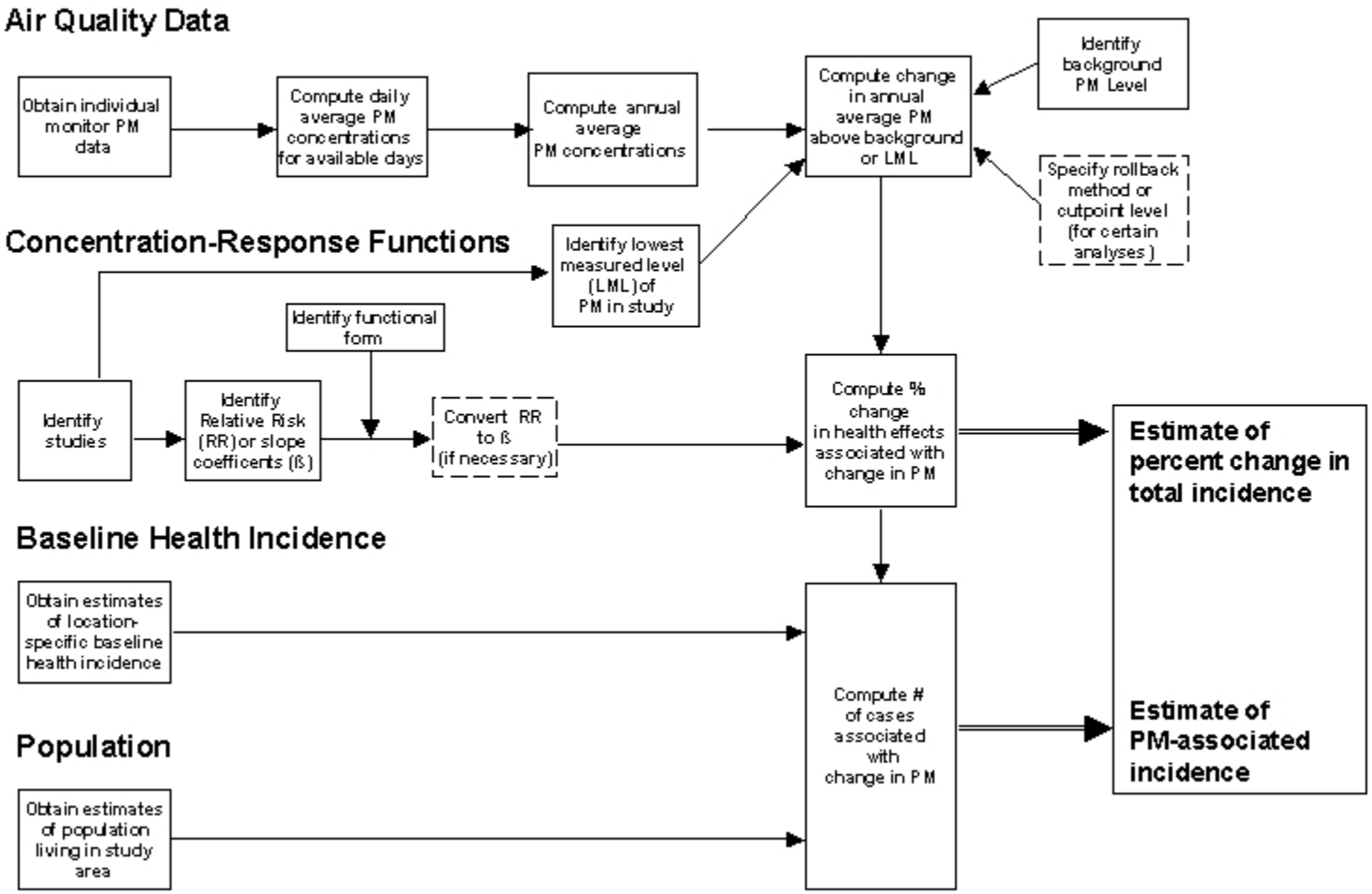


Exhibit 2.3 Flow Diagram of Proposed Risk Analyses for Long-Term Exposure Studies



2.2. Estimating PM background levels

Since health risks will be calculated only for concentrations exceeding estimated background levels, estimates of background PM concentrations in the assessment locations are needed to calculate risk at "as is" concentrations over background and for alternative standard scenarios. The subsections below discuss estimated background levels for $PM_{2.5}$ and $PM_{10-2.5}$, respectively.

2.2.1 Background levels of $PM_{2.5}$

Consistent with the prior PM CD, the draft PM CD (U.S. EPA, 2001a) estimates background annual average $PM_{2.5}$ concentrations in to be in the range of 1 to 4 $\mu g/m^3$ in the Western United States and 2 to 5 $\mu g/m^3$ in the Eastern United States. We propose to use the midpoints of these ranges for the base case analysis and to use the entire ranges in sensitivity analyses. Thus background $PM_{2.5}$ concentrations in the base case analysis are estimated to be 3.5 $\mu g/m^3$ in Boston, Philadelphia, Detroit, and St. Louis; and 2.5 $\mu g/m^3$ in Los Angeles, San Jose, Phoenix, and Seattle.

2.2.2 Background levels of $PM_{10-2.5}$ [to be added later, if appropriate]

2.3 Characterizing "as is" PM air quality

2.3.1 "As is" $PM_{2.5}$ air quality

As discussed earlier, a major input to the $PM_{2.5}$ risk analyses is ambient $PM_{2.5}$ air quality data for each assessment location. In order to be consistent with the approach used in the epidemiological studies that estimated $PM_{2.5}$ C-R functions, the *average* ambient $PM_{2.5}$ concentration on each day for which measured data are available is needed for the risk analyses. Consistent with the approach used in the last PM risk analyses, a composite monitor data set will be created for each assessment location based on a composite of all monitors (except those that are identified as being source-oriented monitors) with at least 11 observations per quarter. At the time of the last PM risk analyses, there was no established $PM_{2.5}$ monitoring network and data sets from special studies conducted in Philadelphia and Los Angeles had to be used. There are now substantial $PM_{2.5}$ air quality data in EPA's Aerometric Information Retrieval System (AIRS) beginning with the year 1999. Where there are sufficient $PM_{2.5}$ data for the year 2000 in AIRS, Abt will use those data for the risk analyses for an assessment location. Alternatively, where there are insufficient air quality data for the year 2000, but sufficient AIRS data from 1999 exist, then Abt will use those data instead.

Appendix A summarizes the $PM_{2.5}$ air quality data that are available for the proposed assessment locations, including quarterly and annual counts, annual averages, and the 98th

percentile of the daily (24-hour) averages. Because the air quality data are not uniformly complete, annual averages will be calculated as the average of quarterly averages to minimize the possible bias resulting from differential amounts of missing data in different quarters of the year.

2.3.2 "As is" $PM_{10-2.5}$ air quality [to be added later, if appropriate]

2.4. Simulating PM levels that meet alternative PM standards

2.4.1 Just meeting alternative $PM_{2.5}$ standards

Predicting the change in risk due to a change in air quality from an "as is" annual mean to meet a lower annual standard when using a C-R function from a long-term exposure study is straightforward: the "as is" mean is simply reduced to the standard level. In this case, simulating just meeting an annual standard does not involve generating an alternative set of daily $PM_{2.5}$ concentrations, because the C-R function estimated in a long-term exposure study is based on annual, rather than daily $PM_{2.5}$ concentrations.

It is more complicated when a C-R function from a short-term exposure study is used. In that case, $PM_{2.5}$ levels that would result from just meeting alternative standards are best modeled by changing the distribution of daily $PM_{2.5}$ concentrations. This section discusses the method proposed for changing daily $PM_{2.5}$ concentrations in an assessment location to simulate just meeting alternative standards.

The form of the $PM_{2.5}$ standards promulgated in 1997 requires that the 3-year average (rounded to the nearest 0.1 $\mu g/m^3$) of the monitor-specific annual means must be at or below the level of the annual standard and the 3-year average (rounded to the nearest 0.1 $\mu g/m^3$) of the ninety-eighth percentile values *at each monitor* cannot exceed the level of the daily standard. In determining attainment of the annual average standard an area may choose to use either the spatially averaged concentrations across all population-oriented monitors or it may use the highest 3-year average based on individual monitors. The most precise simulation of just meeting both the annual and the daily standards in a location would require changing the distribution of daily $PM_{2.5}$ concentrations at each monitor separately. This would require extensive analysis and assumptions about the nature of future control strategies that was considered beyond the scope of the previous risk analyses and is similarly considered beyond the scope of the proposed risk analyses. Therefore, although the amount or percent of reduction on a given day might be determined by the $PM_{2.5}$ concentration at a single monitor on a single day, consistent with the approach used in the prior PM risk analyses, daily $PM_{2.5}$ concentrations will be changed only at a "composite monitor,"whose $PM_{2.5}$ concentration on a given day is the average of the $PM_{2.5}$ concentrations of those monitors reporting on that day. In addition,

although the standard refers to the 3-year average, because of the limited number of years for which $PM_{2.5}$ data are available, we propose to use only a single year of data for the risk analyses.[6]

There are many possible ways to create an alternative distribution of daily concentrations that just meets a specified set of $PM_{2.5}$ standards. The prior PM risk analyses used a proportional (i.e., linear) rollback of all PM concentrations exceeding the estimated background concentration for its base case estimates. This choice was based on analyses of historical $PM_{2.5}$ data which found that year-to-year reductions in $PM_{2.5}$ levels in a given location tended to be roughly linear. That is, both high and low daily $PM_{2.5}$ levels decreased proportionally from one year to the next (see Abt Associates Inc., 1996, Section 8.2). This suggests that, in the absence of detailed air quality modeling, a reasonable method to simulate the $PM_{2.5}$ reductions that would result from just meeting a set of alternative standards would be proportional (linear) rollbacks -- i.e., to decrease $PM_{2.5}$ levels on all days by the same percentage. Appendix B describes a new analysis of historical air quality data for Philadelphia and Los Angeles which continues to support the hypothesis that changes in $PM_{2.5}$ levels that would result if a $PM_{2.5}$ standard were just met are reasonably modeled by using a proportional rollback approach. We recognize that the historic changes in $PM_{2.5}$ have not been the result of a $PM_{2.5}$ control strategy, but likely result from control programs for other pollutants (especially PM_{10}, nitrogen oxides, and sulfur) and from meteorological variability. The pattern of changes that have occurred in the past, therefore, may not necessarily reflect the changes that will result from future efforts to attain $PM_{2.5}$ standards. However, it is interesting to note that reductions in ambient $PM_{2.5}$ concentrations are reasonably modeled by proportional rollbacks in both Philadelphia and Los Angeles, which likely had very different reductions in terms of types of emissions over this period. While there still is uncertainty about the shape of the $PM_{2.5}$ daily air quality distribution that will occur in the future, this may not be a very large uncertainty unless very different control strategies are used in the future (e.g., seasonal controls).

Based on the above considerations, we propose to simulate just meeting the current $PM_{2.5}$ standards and any alternatives that EPA might consider by use of a proportional rollback procedure for the base case estimates. That is, average daily $PM_{2.5}$ concentrations at the composite monitor will be reduced by the same percentage on all days. The percentage reduction will be determined by comparing the maximum of the monitor-specific annual averages (or the maximum of the monitor-specific ninety-eighth percentile daily values) adjusted for background with the level of the annual standard (or daily standard) adjusted for background. Because pollution abatement methods are applied largely to anthropogenic sources of $PM_{2.5}$, rollbacks will be applied only to $PM_{2.5}$ above estimated background levels. The percent reduction will be determined by the controlling standard. For example, suppose both an annual and a daily $PM_{2.5}$ standard are being simulated. Suppose p_a is the percent reduction required to just meet the

[6]The use of a single year of data may be viewed as equivalent to assuming the distribution of $PM_{2.5}$ concentrations is the same for each year during a three-year period.

annual standard, i.e., the percent reduction of daily $PM_{2.5}$ above background necessary to get the highest of the monitor-specific annual averages down to the annual standard. Suppose p_d is the percent reduction required to just meet the daily standard, i.e., the percent reduction of daily $PM_{2.5}$ above background necessary to get the maximum monitor-specific ninety-eighth percentile daily $PM_{2.5}$ value down to the daily standard. If p_d is greater than p_a, then all daily average $PM_{2.5}$ concentrations above background are reduced by p_d percent. If p_a is greater than p_d, then all daily average $PM_{2.5}$ concentrations are reduced by p_a percent.

The proposed method of rollbacks to meet a set of annual and daily $PM_{2.5}$ standards is summarized as follows:

1. Calculate the annual average of $PM_{2.5}$ concentrations at each monitor[7];
2. Calculate the maximum of these monitor-specific annual averages, denoted aa_{max};
3. The percent by which the above-background portion of all daily $PM_{2.5}$ concentrations (at the composite monitor) would have to be reduced to just meet the annual standard (denoted std_a) is

$$p_a = 1 - \frac{(std_a - b)}{(aa - b)}$$

where b denotes background.

4. Calculate the ninety-eighth percentile value of the distribution of daily $PM_{2.5}$ concentrations at each monitor[8];
5. Calculate the maximum of these monitor-specific ninety-eighth percentile values, denoted $98\%ile_{max}$;
6. The percent by which the above-background portion of all daily $PM_{2.5}$ concentrations (at the composite monitor) would have to be reduced to just meet the daily standard (denoted std_d) is

$$p_d = 1 - \frac{(std_d - b)}{(98\%ile_{max} - b)} \quad .$$

[7] Because there are missing air quality data, annual averages will be calculated as the average of quarterly averages to minimize possible bias resulting from differential amounts of missing data in different quarters of the year.

[8] The method of calculating the ninety-eighth percentile value of daily PM concentrations at a monitor will be consistent with the method used by EPA, as described in Appendix N of the July 18, 1997 Federal Register Notice (available on the web at www.epa.gov/ttn/oarpg/t1pfpr.html).

7. Let p_{max} = maximum of (maximum of p_a and p_d) and zero.[9]
8. Then, if PM_o denotes the original PM value on a given day (at the composite monitor), the rolled back PM value on that day, denoted PM_{rb}, is

$$PM_{rb} = b + (PM_o - b) * (1 - p_{max}).$$

The inputs to calculate the percent rollbacks necessary to simulate just meeting annual and daily $PM_{2.5}$ standards of 15 and 65 $\mu g/m^3$, respectively, are given in Exhibit 2.4 for St. Louis to illustrate the approach that will be taken in all locations. The controlling standard (i.e., daily or annual) and corresponding percent rollback necessary to just meet the current $PM_{2.5}$ standards in St. Louis are also shown in Exhibit 2.4. Since an area could potentially use the spatial average of the population-oriented monitors to determine whether or not it met the annual average standard, the risk analysis draft report also will report the percent rollbacks that would have resulted from using this alternative approach in each urban study area.

[9] If the percent rollback necessary to just meet the annual standard and the percent rollback necessary to just meet the daily standard are both negative -- i.e., if both standards are already met -- then the percent rollback applied in the risk analysis is zero. That is, PM values are never increased.

Exhibit 2.4. Inputs to Calculation of Rollbacks to Simulate Just Meeting Annual and Daily $PM_{2.5}$ Standards of 15 $\mu g/m^3$ and 65 $\mu g/m^3$, Respectively, in St. Louis, MO

Monitor Site	Annual Average $PM_{2.5}$ Concentration ($\mu g/m^3$)	98th Percentile $PM_{2.5}$ Concentration ($\mu g/m^3$)
AIRS 171192009881011	16.0	36.3**
AIRS 171634001881011	15.0	32.8
AIRS 290990012881011	14.8	27.4
AIRS 291831002881011	14.9	34.4
AIRS 291892003881011	14.8	30.8
AIRS 291895001881011	14.4	33.3
AIRS 295100085881011	16.4*	34.8
AIRS 295100086881011	15.0	33.2
Maximum of Annual Averages:	**16.4**	---
Percent rollback to just meet an annual standard of 15 $\mu g/m^3$	**10.9%**	**---**
Maximum of 98th Percentile Values:	**---**	**36.3**
Percent rollback to just meet a daily standard of 65 $\mu g/m^3$	**---**	**0**
Controlling standard in St. Louis: Annual Percent rollback required to meet both standards is 10.9%		

*Controlling monitor for the annual standard.

**Controlling monitor for the daily standard.

As noted earlier, proportional (linear) rollback is only one of many possible ways to create an alternative distribution of daily concentrations to meet new $PM_{2.5}$ standards. One could, for example, reduce the high days by one percentage and the low days by another percentage, choosing the percentages so that the new distribution achieves the new standard. At the opposite end of the spectrum from proportional rollbacks, it is possible to meet a daily standard by "peak shaving." The peak shaving method would reduce all daily $PM_{2.5}$ concentrations above a certain concentration to that concentration (e.g., the standard) while leaving daily concentrations at or below this value unchanged. While a strict peak shaving method of attaining a standard is unrealistic, it is illustrative of the principal that patterns different from a proportional rollback might be observed in areas attempting to come into compliance with revised standards. Because the reduction method to attain a daily standard could have a significant impact on the risk analysis results, sensitivity analyses will be conducted on alternative rollback methods reflecting different types of control strategies for meeting the

current or any proposed standards. As with the sensitivity analyses performed for the prior risk analyses, it is proposed that these sensitivity analyses include alternative methods such as reducing the upper 10% of the $PM_{2.5}$ air quality distribution more than the remaining 90% of the distribution.

2.4.2 Just meeting alternative $PM_{10-2.5}$ standards [to be added later, if appropriate]

2.5. Concentration-response functions and estimating health effect incidence changes

The C-R functions considered for use in the planned risk analyses are empirically estimated relations between average ambient concentrations of the pollutant of interest ($PM_{2.5}$ or $PM_{10-2.5}$) and the health endpoints of interest (e.g., short- and long-term exposure mortality or hospital admissions) reported by epidemiological studies for specific locales. This section describes the basic method used to estimate changes in the incidence of a health endpoint associated with changes in PM, using a "generic" C-R function of the most common functional form.

Although some epidemiological studies estimated linear C-R functions, most of the studies used a method referred to as "Poisson regression" to estimate exponential C-R functions in which the natural logarithm of the health endpoint is a linear function of PM[10]:

$$y = B\, e^{\beta x} \quad , \qquad (1)$$

where x is the ambient PM level, y is the incidence of the health endpoint of interest at PM level x, β is the coefficient of ambient PM concentration, and B is the incidence at x=0, i.e., when there is no ambient PM. The relation between a specified ambient PM level, x_0, for example, and the incidence of a given health endpoint associated with that level (denoted as y_0) is then

$$y_0 = B\, e^{\beta x_0} \quad . \qquad (2)$$

Because the exponential form of C-R function (equation (1)) is by far the most common form, the discussion that follows assumes this form.

---

[10]Poisson regression is essentially a linear regression of the natural logarithm of the dependent variable on the independent variable, but with an error structure that accounts for the particular type of heteroskedasticity that is believed to occur in health response data. What matters for the risk analysis, however, is simply the form of the estimated relation, as shown in equation (1).

Ambient PM levels may be based on any averaging time, e.g., they may be daily averages, two-day averages, or annual averages. C-R functions that use as input daily average PM levels relate these to the daily incidence of the health endpoint. There are several variants of the short-term (daily) C-R function. Some C-R functions were estimated by using moving averages of PM to predict daily health effects incidence. Such a function might, for example, relate the incidence of the health effect on day t to the average of PM concentrations on days t and (t-1). Some C-R functions consider the relation between daily incidence and daily average PM lagged a certain number of days. For example, a study might estimate the C-R relation between mortality on day t and average PM on day (t-1). The discussion below does not depend on any particular averaging time or lag time and assumes only that the measure of health effect incidence, y, is consistent with the measure of ambient PM concentration, x.

The change in health effects incidence, $\Delta y = y_0 - y$, from y_0 to the baseline incidence rate, y, corresponding to a given change in ambient PM levels, $\Delta x = x_0 - x$, can be derived by dividing equation (2) by equation (1), which yields:

$$\Delta y = y[e^{\beta \Delta x} - 1] \; . \qquad \textbf{(3)}$$

Alternatively, the change in health effects incidence can be calculated indirectly using relative risk. Relative risk (RR) is a well known measure of the comparative health effects associated with a particular air quality comparison. The risk of mortality at ambient PM level x_0 relative to the risk of mortality at ambient PM level x, for example, may be characterized by the ratio of the two mortality rates: the mortality rate among individuals when the ambient PM level is x_0 and the mortality rate among (otherwise identical) individuals when the ambient PM level is x. This is the RR for mortality associated with the difference between the two ambient PM levels, x_0 and x. Given a C-R function of the form shown in equation (1) and a particular change in ambient PM levels, Δx, the relative risk associated with that change in ambient PM, denoted as $RR_{\Delta x}$, is equal to $e^{\beta \Delta x}$. The change in health effects incidence, Δy, corresponding to a given change in ambient PM levels, Δx, can then be calculated based on this relative risk:

$$\Delta y = y(RR_{\Delta x} - 1) \; . \qquad \textbf{(4)}$$

Equations (3) and (4) are simply alternative ways of expressing the relation between a given change in ambient PM levels, Δx, and the corresponding change in health effects incidence, Δy. These equations are the key equations that combine air quality information, C-R information, and baseline health effects incidence information to estimate ambient PM health risk.

Given a C-R function and air quality data (ambient PM values) from an assessment location, then, the change in the incidence of the health endpoint ($\Delta y = y_0 - y$) corresponding to a change in ambient PM level of $\Delta x = x_0 - x$ can be determined. This can either be done with

equation (3), using the coefficient, β, from a C-R function, or with equation (4), by first calculating the appropriate relative risk from the C-R function.

Because the estimated change in health effect incidence, Δy, depends on the particular change in PM concentrations, Δx, being considered, the choice of PM concentration change considered is important. These changes in PM concentrations are generally reductions from the current levels of PM ("as is" levels) to some alternative, lower level(s).

The daily time-series epidemiological studies estimated C-R functions in which the PM-related incidence on a given day depends only on same-day PM concentration or previous-day PM concentration (or some variant of those, such as a two-day average concentration). Such models necessarily assume that the longer pattern of PM levels preceding the PM concentration on a given day does not affect mortality on that day. To the extent that PM-related mortality on a given day is affected by PM concentrations over a longer period of time, then these models would be mis-specified, and this mis-specification would affect the predictions of daily incidence based on the model. The extent to which longer-term (i.e., weekly, monthly, seasonal, or annual) $PM_{2.5}$ exposures confound the relationship observed in the daily time-series studies is unknown. However, there is some evidence, based on analyses of PM_{10} data, that mortality on a given day is influenced by prior PM exposures up to more than a month before the date of death (Schwartz, 2000). Currently, there is insufficient information to adjust for the impact of longer-term exposure on mortality associated with $PM_{2.5}$ exposures and this is an important uncertainty that should be kept in mind as one considers any results from the short-term exposure $PM_{2.5}$ risk analyses.

The first and second parts of the risk analysis are distinguished primarily by the choice of lower PM level(s). When possible, the choice of lower PM level(s) in the first part of the risk analysis will be the lowest PM concentration observed in the study that estimated the C-R function used in the risk analysis. However, some of the short-term exposure PM studies do not report the lowest observed PM concentration. (For example, some studies instead report the lowest decile or quartile values.) When the lowest observed PM concentration is not reported (or if it is lower than background level), analyses in the first part of the risk analysis will consider the range of "as is" PM concentrations in the assessment location down to background PM concentration in that location. The second part will consider the changes in health effects incidence associated with changes from "as is" PM concentrations to PM concentrations that just meet alternative standards.

In contrast to most short-term exposure studies, long-term exposure studies routinely report the lowest observed annual average PM concentration. Risk analyses that use long-term exposure C-R functions will therefore consider the change from "as is" annual average PM in the assessment location to the lowest annual average PM level observed in the study (or background

level, if that is higher), for the "as is" part of the analysis, or the annual average that would just meet alternative standards for the "risk reduction" part of the analysis.

In both parts of the risk analysis, the ambient PM concentrations to which "as is" ambient PM concentrations are compared are generally lower than or equal to "as is" concentrations. Therefore $\Delta x = x_0 - x$ is negative (or zero), and so the corresponding change in incidence of health effects, Δy, is also negative (or zero). That is, there are fewer cases of any given health effect at lower ambient PM levels. Alternatively, $-\Delta y$ may be interpreted as the health effects attributable to PM concentrations between x_0 and x.

Because different epidemiological studies report different estimated C-R functions for a given health endpoint, predicted changes in health effects incidences depend on the C-R function used. The uncertainty introduced into the risk analyses by this is assessed both through sensitivity analyses and through Monte Carlo methods.

2.6. Calculating health effects incidence on an annual basis

The planned risk analyses will estimate health effects incidence, and changes in incidence, on an annual basis. For mortality, both short-term and long-term exposure studies have reported estimated C-R functions. As noted above, the short-term exposure C-R functions estimated by daily time-series epidemiological studies relate daily mortality to same-day PM concentration or previous-day PM concentration (or some variant of those).

To estimate the *daily* health impacts of daily average ambient PM levels above background or above the levels necessary to just meet a given standard, C-R functions from short-term exposure studies will be used together with estimated changes in daily ambient PM concentrations to calculate the daily changes in the incidence of the health endpoint. (Alternative assumptions about the range of PM levels associated with health effects will be explored in sensitivity analyses. Where a minimum concentration for effects is considered, reductions below this concentration will not contribute attributable cases to the calculation. Only reductions down to this concentration contribute attributable cases to the calculation.)

After daily changes in health effects are calculated, an annual change is calculated by summing the daily changes. However, there are some days for which no ambient PM concentration information is available. The predicted annual risks, based on those days for which air quality data are available, must be adjusted to take into account the full year. If days with missing air quality data occur randomly or relatively uniformly throughout the year, a simple adjustment can be made to the health effect incidence estimate – the incidence estimate based on the set of days with air quality data can be multiplied by the ratio of the total number of days in the year (365) to the number of days in the year for which direct observations were

available, to generate an estimate of the total annual incidence of the health effect.[11] However, if the missing data are not uniformly distributed throughout the year, such a simple adjustment could lead to a biased estimate of the total annual incidence change. To avoid such possible bias, adjustments will be made on a quarterly basis. If Q_i is the total number of days in the ith quarter, and n_i is the number of days in the ith quarter for which there are air quality data, then the predicted incidence change in the ith quarter, based on those days for which there are air quality data, will be multiplied by Q_i/n_i. The adjusted quarterly incidence changes will be summed to derive an estimate of the annual incidence change.

Some short-term exposure C-R functions are based on average PM levels during several days. If these C-R functions are used, the air quality data will be averaged for the same number of days. For example, a function based on two-day averages of PM would be used in conjunction with two-day averages of PM in the assessment location to predict the incidence of the health effect in that location. In some cases, intervals of three or more consecutive days in a given location may be missing data, and so no multi-day average is available for use with multi-day C-R functions. These cases will be treated by multi-day functions just as individual missing days will be treated by single-day functions: they will contribute no incidence change to the risk analysis, and incidence changes will be adjusted for the days on which multi-day averages are missing.

C-R functions from long-term exposure studies (see Exhibit C.4) will be used to assess the annual health impacts of changes in annual average ambient PM concentrations. Once again, to minimize the chance of bias due to differential amounts of missing data in different quarters of the year, quarterly averages will be calculated based on the days in each quarter for which air quality data are available, and the "as is" annual average concentration will then be calculated as an average of the four quarterly averages.

The mortality associated with long-term exposure is likely to include mortality related to short-term exposures as well as mortality not tightly linked to daily changes in PM concentrations. As discussed previously, estimates of daily mortality based on the time-series studies also are likely to be confounded by prior PM exposures. Therefore, the estimated annual incidences of mortality calculated based on the short- and long-term exposure studies should not be added together.

[11] This assumes that the distribution of PM concentrations on those days for which data are missing is essentially the same as the distribution on those days for which we have PM data.

2.7. Baseline health effects incidence data

As noted above, the form of C-R function most commonly used in epidemiological studies on PM, shown in equation (1), is log-linear. To estimate the change in incidence of a health endpoint associated with a given change in PM concentrations using this form of C-R function requires the baseline incidence rate of the health endpoint, that is, the number of cases per unit time (e.g., per year) in the location *before* a change in PM air quality (denoted y in equations 3 and 4).

Incidence rates express the occurrence of a disease or event (e.g., asthma episode, death, hospital admission) in a specific period of time, usually per year. Rates are expressed either as a value per population group (e.g., the number of cases in Philadelphia County) or a value per number of people (e.g., the number of cases per 10,000 residents in Philadelphia County), and may be age and sex specific. Incidence rates vary among geographic areas due to differences in population characteristics (e.g, age distribution) and factors promoting illness (e.g., smoking, air pollution levels).

Incidence rates are available for mortality and for specific communicable diseases which state and local health departments are required to report to the federal government. In addition to the required federal reporting, many state and local health departments collect information on some additional endpoints. These most often are restricted to hospital admission or discharge diagnoses, which are collected to assist in planning medical services. None of the morbidity endpoints in the risk analyses are required to be reported to the federal government.

Although federal agencies collect incidence data on many of the endpoints covered in the PM risk analyses, their data are often available only at the national level, or at the regional or state level. One important exception is mortality rates, which are available at the county level. Because baseline incidence rates can vary from one location to another, location-specific baseline incidence information will be obtained whenever possible. Because hospital admission rates are available for some locations and not others, this was a consideration in the selection of locations for which to conduct the PM risk analyses. For respiratory symptom or illness health endpoints, the only estimates of baseline incidence rates available are typically from the studies that estimated the C-R functions for those endpoints. However, because risk analysis locations for these endpoints were selected partly on the basis of where studies were carried out, baseline incidence rates reported in the studies should be appropriate to the risk analysis locations to which they are applied. A more detailed discussion of baseline health effects incidence data is presented in Section 5.

2.8. Characterizing uncertainty

Any estimation of "as is" risk and risk reductions under current or alternative standard scenarios will involve substantial uncertainties, and there are additional uncertainties for a pollutant such as PM (as opposed to, for example, ozone), given the diversity of composition in this generally defined pollutant. The following will be among the major sources of uncertainty in the risk analyses:

- Uncertainties related to estimating the C-R functions including the following:
  - There is statistical uncertainty surrounding estimates of $PM_{2.5}$ (and $PM_{10-2.5}$) coefficients in C-R functions used in the analyses.
  - There is uncertainty about the shape of the C-R relationship, particularly whether or not there are thresholds below which no response occurs.
  - There is uncertainty related to the transferability of PM C-R functions from study locations to the locations selected for the risk analyses.[12] A C-R function in a study location may not provide an accurate representation of the C-R relationship in the analysis location(s) because of
    - variations in PM composition across cities,
    - the possible role of associated co-pollutants in influencing PM risk,
    - variations in the relation of total ambient exposure (both outdoor and ambient contributions to indoor exposure) to ambient monitoring in different locations (e.g, due to differences in air conditioning use in different regions of the U.S.),
    - differences in population characteristics (e.g., the proportions of members of sensitive subpopulations) and population behavior patterns across locations.

- Uncertainties related to the air quality adjustment procedure that will be used to simulate just meeting current or alternative PM standards, and uncertainties about estimated background concentrations for each location.

- Uncertainties associated with use of baseline health effects incidence information that is not specific to the analysis locations.[13]

---

[12] The proposed risk analysis locations were selected partly on the basis of where C-R functions were estimated, specifically to avoid this important source of uncertainty. Therefore, this will be a source of uncertainty in the risk analyses only when C-R functions from multi-city studies or from another location are applied to a risk analysis location.

[13] This is not an issue for mortality, because county-specific mortality rates are available. Because proposed risk analysis locations have been selected partly on the basis of where C-R functions were estimated, this also will not be a source of substantial uncertainty for risk analyses of respiratory symptoms and illnesses. The studies that

The uncertainties from some of these sources -- in particular, the statistical uncertainty surrounding estimates of the PM coefficients in C-R functions -- can be characterized quantitatively. It will be possible, for example, to calculate confidence intervals around risk estimates based on the uncertainty associated with the estimates of pollutant coefficients used in the risk analyses. These confidence intervals will express the range within which the true risks are likely to fall *if the uncertainty surrounding PM coefficient estimates were the only uncertainty in the analysis*. There are, of course, several other uncertainties in the risk analysis, as noted above. If there were sufficient information to quantitatively characterize these sources of uncertainty, they could be included in a Monte Carlo analysis to produce confidence intervals that more accurately reflect all sources of uncertainty.

There are several ways to handle uncertainties in the risk analyses:

- Limitations and assumptions in estimating risks and risk reductions will be clearly stated and explained.

- For any endpoint for which only a single C-R function has been estimated, the uncertainty resulting from the statistical uncertainty associated with the estimate of the pollutant coefficient will be characterized by confidence intervals around the point estimate of risk. As noted above, such a confidence interval will express the range within which the true risk is likely to fall *if the uncertainty surrounding the pollutant coefficient estimate were the only uncertainty in the analysis*. It will not, for example, reflect the uncertainty concerning whether the pollutant coefficients in the study location and the assessment location are the same.[14]

- Sensitivity analyses will be conducted to illustrate the effects of changing key default assumptions on the results of the assessment, and quantitative comparisons will be presented to inform other analytic choices.[15]

Possible additional or alternative approaches to characterizing uncertainty that are being considered include the following:

---

estimated C-R functions for these health endpoints generally reported baseline incidence rates. Because the proposed risk analysis locations are in locations where C-R functions have been estimated, the studies provide baseline incidence rates that are appropriate to the risk analysis locations.

[14] This is not an uncertainty, of course, if the C-R function has been estimated in the assessment location.

[15] "Sensitivity analyses" refers to assessing the effects of uncertainty on some of the final risk estimates; "quantitative comparisons" refer to numerical comparisons (e.g. comparisons of monitor values) that are not carried that far.

- "integrated sensitivity analyses" may be presented to include in an overall assessment of uncertainty those sources of uncertainty that cannot readily be quantified. Such analyses would rely on informed judgments to assign probabilities to possible alternatives. For example, judgment could be made concerning the likelihood that each of several possible alternative assumptions is the correct one. This procedure allows sources of uncertainty that otherwise cannot be quantified to be included in a Monte Carlo sensitivity analysis.

- Different sets of plausible assumptions that would result in "low end," "middle," and "high end" estimates of incidence could be identified, and the estimates resulting under each set of assumptions could be presented as alternatives.

2.9. Sensitivity analyses

Sensitivity analyses can be used to illustrate the sensitivity of analysis results to different possible input values or to different assumptions or procedures that may affect these input values. Although a sensitivity analysis is not as comprehensive as an uncertainty analysis, selecting only a few possible alternative values of an input component rather than characterizing the entire distribution of these values, it is precisely the simplicity of a sensitivity analysis that makes it preferable for illustrating the impact on results of using different input component values. Exhibit 2.5 lists the proposed sensitivity analyses.

Exhibit 2.5. Planned Sensitivity Analyses and Quantitative Comparisons

Analysis Number (Exhibit 2. 1)	Component of the Risk Analysis	Sensitivity Analysis or Comparison
1	Air Quality	A sensitivity analysis of the effect of different assumptions about background PM levels
2	Air Quality	A sensitivity analysis of the effect of different air quality adjustment procedures on the estimated risk reductions resulting from just meeting alternative 24-hr and annual standards
3	Baseline Incidence	A comparison of using more aggregate incidence data (national, state, etc) versus county-specific information in the county with the best local incidence data
4	Concentration-Response	A comparison or sensitivity analysis using an approach to estimate the possible impact of using a distributed lag C-R function.
5	Concentration-Response	A comparison or sensitivity analysis of the impact on mortality associated with long-term exposure of different assumptions about the role of historical air quality concentrations in contributing to the reported effects.
6	Concentration-Response	A sensitivity analysis using C-R functions for PM from multi-pollutant regressions with co-pollutants versus single pollutant regressions
7	Concentration-Response	A sensitivity analysis assuming alternative potential threshold concentration levels for the occurrence of PM-related response at concentrations above those for background

3. Selecting Health Endpoints and Urban Areas

For the prior PM NAAQS review there was a fairly limited number of studies that directly measured $PM_{2.5}$ and which were judged suitable for use in the $PM_{2.5}$ risk analyses. As discussed in the 2001 draft PM CD, a significant number of epidemiological studies examining a variety of health effects associated with ambient $PM_{2.5}$ concentrations in various locations throughout the United States and Canada have been published since the last review. A smaller subset of these studies also have examined the relationship between ambient $PM_{10-2.5}$ concentrations and various health effects. As a result of the availability of additional health effects studies and air quality information, EPA proposes to expand somewhat the geographic scope of the PM risk analyses to include several additional urban areas beyond the two (Philadelphia and Los Angeles Counties) analyzed for the last review, consistent with the goals of the assessment.

Proposed approaches to selection of both health endpoint categories and urban areas to include in the $PM_{2.5}$ risk analyses are discussed below. Similar approaches will be used, if warranted, for possible $PM_{10-2.5}$ risk analyses based on review of the next draft of the PM CD.

3.1. Health endpoints

3.1.1. Health endpoints for $PM_{2.5}$

OAQPS staff has carefully reviewed the evidence evaluated in the 2001 draft PM CD. Tables 9-3 and 9-6 in the 2001draft PM CD summarize the available U.S. and Canadian studies that provide effect estimates for $PM_{2.5}$ and other fine particle indicators for short- and long-term exposures, respectively. Given the large number of endpoints and studies addressing $PM_{2.5}$ effects, EPA is proposing to include in the quantitative risk analyses only the more severe and better understood (in terms of health consequences) health endpoint categories for which the weight of the evidence supports the existence of a relationship between $PM_{2.5}$ and the effect category. In addition, only those categories which included studies that directly measured fine fraction PM using $PM_{2.5}$ or $PM_{2.1}$ as the indicator are proposed to be included. Based on its review of the evidence evaluated in the 2001 draft PM CD, OAQPS proposes to include the following broad categories of health endpoints in the $PM_{2.5}$ analyses:

- non-accidental total, cardiovascular and respiratory mortality (due to short-term exposure)
- total mortality (due to long-term exposure)
- hospital admissions for cardiovascular and respiratory causes (due to short-term exposure)
- emergency room visits for cardiovascular and respiratory causes (due to short-term exposure)

- respiratory illnesses and/or symptoms not requiring hospitalization (due to short-term exposure)

Other effects reported to be associated with $PM_{2.5,}$ such as decreased lung function and changes in heart-rate variability will be addressed qualitatively in the OAQPS PM Staff Paper.

3.1.2. Health endpoints for $PM_{10-2.5}$ [to be added, if appropriate]

3.2. Urban areas

In the prior risk analyses the selection of urban areas to include in the analyses was largely determined by the very limited availability of recent and sufficiently complete $PM_{2.5}$ ambient air quality data. For this review, there are a significantly greater number of candidate locations, at least for the $PM_{2.5}$ analyses, in which epidemiological studies have reported C-R relationships and for which there are sufficient $PM_{2.5}$ ambient air quality data. Recent evidence from the National Mortality and Morbidity Air Pollution Study (NMMAPS) (Samet et al., 2001) suggests there is geographic variability in C-R relationships across many U.S. urban areas. In light of the evidence from NMMAPS, which examined C-R relationships across the 90 largest U.S. cities using PM_{10} as the indicator, we believe it is desirable to conduct the proposed PM risk analyses, to the extent possible, in the urban areas in which C-R relationships have been estimated.

Developing a list of proposed urban areas to include in the risk analyses from the larger candidate pool has been guided by four overriding considerations:

- To the extent possible, urban locations should be the same as or close to the study locations where C-R functions have been estimated for the health endpoints recommended above.[16]
- Uncertainties surrounding estimates of risk and risk reduction should be minimized to the extent possible by focusing on locations in which studies that had greater precision were conducted, as indicated by greater statistical power to detect relatively small population effects.
- The urban areas selected should have recent and sufficient $PM_{2.5}$ and/or $PM_{10-2.5}$ ambient air quality data to support a risk analysis.
- For the hospital admission effects category, the availability of relatively recent baseline incidence data, specific to International Classification of Disease (ICD) codes is important.

---

[16] Urban locations for which C-R functions were estimated often include several counties. (For example, in Schwartz et al., 1996, the urban area labeled "Boston" consisted of three counties: Middlesex, Norfolk, and Suffolk counties.) To the extent possible, in the PM risk analyses we will try to include the specific counties used in the urban location in the original epidemiological studies.

To the extent feasible, it also is desirable to conduct the analyses in the same set of urban areas for the various health endpoint categories and PM indicators.

3.2.1. Urban areas for the $PM_{2.5}$ analyses

The largest data base for health effects associated with short-term (i.e., 24-hour) ambient $PM_{2.5}$ concentrations, in terms of number of studies in different locations, is for non-accidental total and cause-specific mortality. Therefore, OAQPS has focused on selecting urban areas for the risk analyses based mainly, but not exclusively, on this health effect category. Because baseline mortality incidence data are available at the county level, this is not a limiting factor in the selection of urban areas for the $PM_{2.5}$ risk analyses.

Exhibits C.1 and C.2 in Appendix C present a summary of the U.S. and Canadian studies identified in the 2001 draft PM CD that report effect estimates for short-term exposure mortality and (cardiovascular and respiratory) morbidity, respectively, associated with $PM_{2.5}$ or $PM_{2.1}$. The U.S. locations in this exhibit represent the candidate pool of possible locations to include in the risk analyses. The considerations listed above were used in deriving the list of proposed urban areas to be included in the PM risk analyses examining short-term exposure mortality. In narrowing the list from the larger candidate pool, we first considered the statistical power of the studies that estimated $PM_{2.5}$ short-term exposure mortality C-R functions in those locations. In general, the power of a study increases as the number of its observations increases. The number of observations depends not only on the number of days on which mortality counts were obtained, but also on the size of the mortality counts. The 2001 draft PM CD uses the natural logarithm of the mortality-days (i.e., the natural log of the product of the number of study days and the average number of deaths per day) as a surrogate or indicator reflecting the power of short-term exposure mortality epidemiological studies. Exhibit C.1 summarizes the natural log of mortality-days for all of the available $PM_{2.5}$ mortality studies. OAQPS proposes to consider for inclusion in the risk analyses only those locations in which studies with relatively greater statistical power were conducted – specifically, studies that have a natural log of mortality-days greater than or equal to 9.0 for total non-accidental mortality.[17]

EPA next considered which of those study locations also have sufficient $PM_{2.5}$ monitoring data to support a risk analysis. The studies in bold typeface in Exhibit C.1 indicate the non-accidental short-term exposure mortality study locations which had sufficient $PM_{2.5}$ air quality data and in which studies with relatively greater statistical power were conducted. Exhibit C.3 shows the monitor-specific minimum number of observations per quarter and the

[17]Most of the epidemiological studies reporting total non-accidental mortality, also report on one or more cause specific mortality categories; in such studies the natural log of mortality days is often less than 9.0 because there are fewer deaths from a specific cause. We plan to include the cause-specific mortality C-R relationships reported in such studies as long as the natural log of total mortality days is greater than or equal to 9.0.

number of observations per year for all of the U.S. locations for the studies summarized in Exhibits C.1 and C.2. A city was considered to have sufficiently complete air quality data if it had at least one monitor at which there were at least 11 observations per quarter and at least 122 observations per year (i.e., equivalent to at least 1 in 3 day monitoring).[18] Using that completeness criterion, six cities in which epidemiological studies reported C-R relationships for $PM_{2.5}$ and mortality and which had sufficient year 2000 data are listed in bold in Exhibit C.3. As indicated in Exhibit C.3, although Phoenix did not have sufficient data in 2000, it did have sufficient data in 1999.

We propose to exclude those monitors which are identified in AIRS as targeting "highest concentration" as their monitoring objective, which are generally located in either an "industrial" or "commercial" land use area based on the information from AIRS. For any monitor which is not thus excluded but is not specifically identified as "population-oriented," Abt plans to evaluate the extent to which the daily $PM_{2.5}$ concentrations at the monitor are correlated with those at the other monitors in the urban area to which it belongs. In cases where there is a very low correlation, we propose to drop the monitor from the analysis and to exclude the population living near it.

Based on the criteria of study power and availability of sufficiently recent and complete air quality data, OAQPS proposes to include the following urban areas in a $PM_{2.5}$ risk analysis for short-term exposure mortality:

- Philadelphia
- Los Angeles
- Phoenix
- San Jose
- Boston
- Detroit
- St. Louis

The long-term exposure C-R functions proposed to be used in the PM risk analyses are based on studies involving multiple cities across the United States (see Exhibit C.4), and the estimated C-R functions are based on differences in long-term averages observed across the various cities. The issue of matching a risk analysis location with the specific location in which a C-R function was estimated, to minimize the uncertainties associated with geographic differences, therefore does not arise for long-term exposure mortality in quite the way it does for short-term exposure mortality. We propose to carry out the risk analysis for long-term exposure

[18]To be consistent with the epidemiological studies which generally focus on using only population-oriented monitors, we excluded from consideration any monitors where the monitoring objective was listed as "highest concentration monitor." The few monitors that were excluded were sited in industrial or commercial areas and are intended to characterize local conditions near major point sources.

mortality in the same seven urban locations that were proposed for the short-term exposure mortality risk analyses.

EPA considered the alternative approach of carrying out a national scale analysis for mortality associated with long-term exposure to $PM_{2.5}$. However, EPA does not believe that a national scale analysis is useful or necessary for purposes of the current review of the $PM_{2.5}$ standards. Given the many sources of uncertainty inherent in conducting risk analyses for $PM_{2.5}$, as well as geographical variation in the composition of ambient PM, EPA believes that extrapolating the available data to provide national-scale estimates would introduce large uncertainties in any estimates. EPA staff recognize the limited role of the risk analyses in this standards review and do not plan to use the risk estimates as a principal basis for recommending selection among alternative standard levels. EPA believes focusing its analytical effort on a few selected urban areas where there are adequate PM air quality data and where various sensitivity analyses can be carried out, will be more useful as a means of informing the standards review than an effort to develop national mortality estimates associated with PM exposure.

Most of the urban locations in which C-R functions were estimated for health endpoints other than mortality are included in the set of locations available for mortality (see Exhibit C.2 on hospital admission, emergency room visit, and respiratory symptom and illness studies). A primary consideration in selecting urban locations for these other health endpoints, as with the risk analyses for mortality, is that the assessment locations be the same as or close to the study locations where C-R functions were estimated. Second, studies with sufficient statistical power to detect relatively small but real population effects are preferable. As with mortality, another consideration is the availability of recent and adequate $PM_{2.5}$ air quality data. Finally, for the hospital admission effect category, the availability of baseline incidence data is an additional consideration in selecting urban locations for the risk analyses. Data on hospital admissions for recent years, specific to International Classification of Disease (ICD) codes, are available in some cities but not others. Based on all of the above considerations, the proposed locations for conducting the $PM_{2.5}$ risk analyses for hospital admissions and emergency visits were selected and they are indicated in bold typeface in Exhibit C.2. In addition, we propose to use the estimated C-R relationships reported in Schwartz and Neas (2000), a study conducted across several cities, to estimate risks in Boston and St. Louis for respiratory symptom endpoints.

Based on applying the criteria and considerations discussed above, Exhibit 3.1 displays the recommended study locations and associated health endpoint categories for inclusion in the risk analyses for $PM_{2.5}$.

3.2.2. Urban areas for the $PM_{10-2.5}$ analyses [to be added, if appropriate]

Exhibit 3.1. Proposed Study Locations for $PM_{2.5}$ Concentration-Response Functions, by Endpoint

Health Endpoint		Urban Locations							
		Boston, MA	Detroit, MI	Los Angeles, CA	Philadelphia PA	Phoenix, AZ	San Jose, CA	St. Louis, MO	Seattle, WA
Non-Accidental Short-term Mortality	Total	X	X	X	X	X	X	X	
	Cardiovascular	X	X	X	X	X	X	X	
	Respiratory	X	X	X	X	X	X	X	
Non-Accidental Long-term Mortality	Total	X							
Hospital Admissions & Emergency Room Visits	Cardiovascular		X	X					
	Respiratory		X	X					X
Respiratory Symptoms*		X						X	

*A single C-R function was estimated based on Schwartz and Neas (2000) based on combined data from six urban locations.

4. Selecting Concentration-Response Functions

For the most part the selection of a proposed list of studies from which to draw C-R relationships for the PM risk analyses has already been determined by the choice of health endpoints to include in the analyses and by the process used to develop the proposed list of urban areas to include in the analyses which was discussed in the previous section. In addition to those studies identified in the previous section as providing appropriate C-R relationships to use in the analyses, as discussed below we also are considering the use of C-R functions from additional multi-city and single city studies that were conducted in Canada.

The C-R functions of interest for the PM risk analyses are from epidemiological studies investigating the relations between PM and a variety of health endpoints. C-R functions proposed for possible use in the PM risk analyses have been selected from among those listed in Tables 9-3, 9-4, and 9-6 of the 2001 draft PM CD and include studies that were used in the prior (1996) PM risk analysis (Abt Associates Inc, 1996). As noted earlier, the selection of studies is preliminary and will be reviewed once the next draft PM CD is available.

As can be seen in Exhibits C.1 and C.2 in Appendix C, there are often several possible C-R functions that can be used for a given health endpoint. For some health endpoints there are both single-city studies and multi-city studies. In addition, studies often report more than one estimated C-R function for the same location and health endpoint. Sometimes models including different sets of co-pollutants are estimated in a study; sometimes different lags are estimated. It is also possible that two different studies estimated a C-R function for the same combination of PM and health endpoint in the same location. It is therefore necessary to make decisions about which C-R functions to use in the risk analysis.

4.1. Single and multi-city functions

All else being equal, a C-R function estimated in the assessment location is preferable to a function estimated elsewhere since it avoids uncertainties related to potential differences due to geographic location. That is why the urban areas considered as candidates to be included in the risk analysis were those locations in which C-R functions have been estimated. There are several advantages, however, to using estimates from multi-city studies versus studies carried out in single cities. Multi-city studies are applicable to a variety of settings, since they estimate a central tendency across multiple locations. They also tend to have more statistical power versus single city studies due to larger sample sizes, reducing the uncertainty around the estimated coefficient. Because single-city and multi-city studies have different advantages, if a single-city C-R function has been estimated in a risk analysis location and a multi-city study is also available for the same health endpoint, the results from both will be used and reported in the risk analyses.

4.2. Single and multi-pollutant models

For some of the epidemiological studies identified for obtaining C-R relationships for the risk analyses, C-R functions are reported both for the case where only PM levels were entered into the health effects model (i.e., single pollutant models) and where PM and one or more other measured gaseous co-pollutants (i.e., ozone, nitrogen dioxide, sulfur dioxide, carbon monoxide) were entered into the health effects model (i.e., multi-pollutant models). To the extent that any of the co-pollutants present in the ambient air may have contributed to the health effects attributed to PM in single pollutant models, risks attributed to PM might be overestimated where C-R functions are based on single pollutant models. However, as shown in the preliminary draft PM SP (see Figure 3-11, p.3-62 - 3-63), the magnitude and statistical significance of the associations reported between $PM_{2.5}$ and mortality due to short-term exposure show no trends with the levels of any of the four gaseous co-pollutants examined. As stated in the preliminary draft PM SP, "While not definitive, these consistent patterns indicate that it is more likely that there is an independent effect of $PM_{2.5}$, ... that is not confounded or appreciably modified by the gaseous co-pollutants."(draft PM SP, p.3-64)

For some of the gaseous co-pollutants, such as carbon monoxide, nitrogen dioxide, and sulfur dioxide, which tend to be highly correlated with ambient PM concentrations in some cities, it is difficult to sort out whether these pollutants are exerting any independent effect from that attributed to PM. As discussed in the 2001 draft PM CD, inclusion of pollutants that are highly correlated with one another can lead to misleading conclusions in identifying a specific causal pollutant. When collinearity exists, multi-pollutant models would be expected to produce unstable and statistically insignificant effects estimates for both PM and the co-pollutants (2001 draft PM CD, p.9-81).

Given the lack of consensus on whether single or multi-pollutant models provide more reliable C-R relationships for estimating risks associated with ambient $PM_{2.5}$ concentrations, we propose to report risk estimates based on both single and multi-pollutant models where both are available.

4.3. Single, multiple, and distributed lag functions

There is recent evidence (Schwartz., 2000), that the relation between PM and health effects may best be described by a distributed lag (i.e., the incidence of the health effect on day n is influenced by PM concentrations on day n, day n-1, day n-2 and so on). If this is the case, a model that includes only a single lag (e.g., a 0-day lag or a 1-day lag) is likely to understate the total impact of PM. Because of this, when a study reports several estimated lag models, the one that produces the greatest relative risk is likely to minimize the degree of understatement of models that include only one lag at a time. Therefore, if several lag models have been estimated, we propose to use the model that results in the greatest predicted RR. We also plan to conduct a

sensitivity analysis examining the potential impact of using a distributed lag approach for short-term exposure mortality associated with $PM_{2.5}$, based on the distributed lag analysis of PM_{10} and short-term exposure mortality by Schwartz (2000).

4.4. Summary

To summarize, the basic proposed approach to selecting C-R functions is as follows:

- if a single-city C-R function has been estimated in a risk analysis location and a multi-city study is also available, risk and risk reduction estimates based on both will be reported;

- if both single-pollutant and multi-pollutant C-R functions are available, risk and risk reduction estimates based on both will be reported;

- if several lag models have been estimated, the model that results in the greatest predicted RR will be used. (A sensitivity analysis examining the potential impact of a distributed lag approach also is planned for short-term exposure $PM_{2.5}$ mortality).

5. Baseline Health Effects Incidence Rates

Many of the epidemiology studies proposed for use in the PM risk analyses directly estimate the percentage change in incidence (i.e., the relative risk), rather than the absolute number of cases for an endpoint. To estimate the annual number of PM-associated cases using these studies, it is necessary to know the annual baseline incidence, that is, the annual number of cases in a location *before* a change in PM air quality.

Incidence rates express the occurrence of a disease or event (e.g., asthma episode, death, hospital admission) in a specific period of time, usually per year. Rates are expressed either as a value per population group (e.g., the number of cases in Philadelphia County) or a value per number of people (e.g., number of cases per 10,000 residents), and may be age and sex specific. Incidence rates vary among geographic areas due to differences in population characteristics (e.g, age distribution) and factors promoting illness (e.g., smoking, air pollution levels).[19] The sizes of the populations in the proposed assessment locations that are relevant to the proposed risk analyses (i.e., the populations for which the $PM_{2.5}$ C-R functions are estimated and the baseline incidences refer) are given in Exhibit 5.1. If there is sufficient information to carry out a $PM_{10\text{-}2.5}$ risk analysis, the urban areas that would be selected are likely to be a subset of the locations proposed for the $PM_{2.5}$ risk analysis.

5.1. Sources of incidence data for the $PM_{2.5}$ risk analyses

Incidence rates are available for mortality (death rates) and for specific communicable diseases which state and local health departments are required to report to the federal government. None of the morbidity endpoints proposed for the risk analysis are required to be reported to the federal government. In addition to the required federal reporting, many state and local health departments collect information on some additional endpoints. These most often are restricted to hospital admission or discharge diagnoses, which are collected to assist in planning medical services. Data may also be collected for particular studies of health issues of concern.

Although federal agencies collect incidence data on many of the endpoints proposed to be included in the risk analyses, their data are often available only at the national level (national averages), or at the regional or state level. When possible, Abt contacted state and local health departments and hospital planning commissions to obtain location-specific rates.

[19] Incidence rates also vary within a geographic area due to the same factors; however, statistics regarding within-city variations are rarely available and are not necessary for this analysis.

Exhibit 5.1. Relevant Population Sizes for Proposed Locations.

Population[a]	Philadelphia[1]	Philadelphia Region[2]	Los Angeles[3]	Phoenix[4]	San Jose[5]	Boston[6]	Detroit[7]	St. Louis[8]	Seattle[9]
Total	1,518,000	4,603,000	9,519,000	3,072,000	1,683,000	2,806,000	2,061,000	2,518,000	1,737,000
Ages ≥25	966,000 (64%)	3,031,000 (66%)	5,871,000 (62%)	1,931,000 (63%)	1,110,000 (66%)	1,903,000 (68%)	1,303,000 (63%)	1,637,000 (65%)	---
Ages ≥30	852,000 (56%)	2,733,000 (59%)	5,092,000 (53%)	1,684,000 (55%)	965,000 (57%)	1,673,000 (60%)	1,153,000 (56%)	1,475,000 (59%)	---
Ages 0-19	---	---	2,947,000 (31%)	---	---	---	---	---	---
Ages 20-64	---	---	5,646,000 (59%)	---	---	---	---	---	---
Ages ≥ 65	---	---	927,000 (10%)	359,000 (12%)	---	---	---	---	---
Children, ages 7-14	---	---	---	---	---	283,000 (10%)	---	307,000 (12%)	---

[a] Total population and age-specific population estimates are based on 2000 U.S. Census data. See http://factfinder.census.gov/. Populations are rounded to the nearest thousand. The urban areas given in this exhibit are those considered in the studies proposed to be used in the PM risk analyses. Lipfert et al. (2000) used Philadelphia County when estimating the C-R function for non-accidental mortality and a larger collection of counties, denoted in this exhibit as Philadelphia Region, when estimating C-R functions for cardiovascular and respiratory mortality.

[1] Philadelphia County.
[2] Philadelphia, Bucks, Delaware, Montgomery, Camden (NJ), Gloucester (NJ), and Burlington Counties (NJ).
[3] Los Angeles County.
[4] Maricopa County.
[5] Santa Clara County.
[6] Middlesex, Norfolk, and Suffolk Counties.
[7] Wayne County.
[8] St. Louis, Franklin, Jefferson, St. Charles, Clinton (IL), Madison (IL), Monroe (IL), and St. Clair (IL) Counties and St. Louis City.
[9] King County.

Abt obtained estimates of location-specific baseline mortality rates for each of the proposed assessment locations for 1998 from CDC Wonder, an interface for public health data dissemination from the Centers for Disease Control (CDC).[20] The mortality rates are derived from U.S. death records and U.S. Census Bureau post-censal population estimates, and are reported in Exhibit 5.2 per 100,000 general population. In all cases, the incidence rates listed correspond to the ages of the populations studied in the relevant epidemiology studies, e.g., individuals over 65 years of age. National rates are provided for comparison for 1998 from CDC Wonder. Mortality rates were not obtained for King County, Washington because none of the proposed mortality C-R functions were estimated in this location.

Baseline incidence rates for both cardiovascular and respiratory hospital admissions will be obtained, if available, for those locations in which hospitalization C-R functions were estimated: Detroit, Los Angeles, and Seattle. Hospitalization data for Los Angeles County in 1999 have been obtained from California's Office of Statewide Health Planning and Development Data Users Support Group. This data are presented in Exhibit 5.3 for the hospitalization endpoints associated with the C-R functions estimated in Los Angeles. Abt is in the process of obtaining hospitalization rates for Seattle and is looking into obtaining similar data for the remaining locations. For Los Angeles, the data are actually annual hospital discharge data, which will be used as a proxy for hospital admissions. By using the annual discharge rate, we assume that the admissions at the end of the year 1999 that carry over to the beginning of 2000 (and are therefore not included in the 1999 discharge data) are offset by the admissions in 1998 that carry over to the beginning of 1999 (and are therefore included in the 1999 discharge data) for each condition.

Baseline incidence rates for emergency room (ER) visits are less likely to be readily accessible than location-specific hospitalization data. In order to estimate ER visits in each of the assessment locations, the national ratio of ER visits to hospital discharges will be determined for each of the relevant health conditions. These data will be obtained from the National Center for Health Statistics (NCHS).[21] The ratio of the number of nationwide hospital admissions to ER visits for a given condition will be applied to the hospitalization rates estimated in each assessment location to derive the baseline rate of ER visits. This approach assumes that there is no differential utilization of the emergency room relative to hospitalizations between the assessment locations and the nation as a whole.

---

[20] See http://wonder.cdc.gov/.

[21] NCHS has conducted two nationwide surveys: the 1999 National Hospital Ambulatory Medical Care Survey (which can be downloaded from: ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Datasets/NHDS/) and the 1999 National Hospital Discharge Survey (which can be downloaded from: ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Datasets/NHAMCS/.

Exhibit 5.2. Baseline Mortality Rates for 1998 for Proposed $PM_{2.5}$ Risk Analysis Locations.

Health Effect	Philadelphia[1]	Philadelphia Region[2]	Los Angeles[3]	Phoenix[4]	San Jose[5]	Boston[6]	Detroit[7]	St. Louis[8]	National Average[a]
Mortality:									
A. Short-term Exposure Mortality[b] (per 100,000 general population/year)									
Non-accidental (all ages): ICD codes <800	1137	---	600	690	518	797	891	909	808
Non-accidental (65+): ICD codes <800	---	---	---	536	---	---	---	---	633
Cardiovascular (all ages): ICD codes: 390-459	---	---	---	---	221	---	406	---	349
Cardiovascular (all ages): ICD codes: 390-448	---	386	---	277	---	---	---	---	348
Cardiovascular (all ages): ICD codes: 390-429	---	---	219	---	---	---	---	---	273

Health Effect	Philadelphia[1]	Philadelphia Region[2]	Los Angeles[3]	Phoenix[4]	San Jose[5]	Boston[6]	Detroit[7]	St. Louis[8]	National Average[a]
Respiratory (all ages): ICD codes: 11, 35, 472-519, 710.0, 710.2, 710.4	---	---	---	---	68	---	---	---	89
Respiratory (all ages): ICD codes: 460-519	---	95	---	---	---	---	80	---	88
COPD and Asthma (all ages): ICD codes: 490-496	---	---	31	---	---	---	---	---	42
B. Long-term Exposure Mortality[a] (per 100,000 general population/year)									
Total mortality (25+): ICD codes: all	1179	977	618	725	533	819	920	938	837
Total mortality (30+): ICD codes: all	1165	969	610	689	527	813	907	929	830
Cardiopulmonary Mortality (25+): ICD codes: 400-440, 485-495	484	413	311	302	249	338	430	437	370

Health Effect	Philadelphia[1]	Philadelphia Region[2]	Los Angeles[3]	Phoenix[4]	San Jose[5]	Boston[6]	Detroit[7]	St. Louis[8]	National Average[a]
Cardiopulmonary Mortality (30+): ICD codes: 401-440, 460-519	543	464	340	348	276	385	466	486	415

a. Mortality figures were obtained from CDC Wonder for 1998. See http://wonder.cdc.gov/.

b. Mortality rates are presented only for the locations in which the C-R functions were estimated. Lipfert et al. (2000) used Philadelphia County when estimating the C-R function for non-accidental mortality and a larger collection of counties, denoted in this exhibit as Philadelphia Region, when estimating C-R functions for cardiovascular and respiratory mortality. See Exhibit C.1 in Appendix C for $PM_{2.5}$ C-R functions proposed to be used in the analyses. All incidence rates are rounded to the nearest unit. Mortality rates for St. Louis may be slightly underestimated because some of the mortality counts in the smaller counties were reported as missing in CDC Wonder. We are currently examining this issue further.

[1] Philadelphia County. [2] Philadelphia, Bucks, Delaware, Montgomery, Camden (NJ), Gloucester (NJ), and Burlington Counties (NJ). [3] Los Angeles County. [4] Maricopa County. [5] Santa Clara County. [6] Middlesex, Norfolk, and Suffolk Counties. [7] Wayne County. [8] St. Louis, Franklin, Jefferson, St. Charles, Clinton (IL), Madison (IL), Monroe (IL), and St. Clair (IL) Counties and St. Louis City. [9] King County.

Exhibit 5.3. Baseline Hospitalization Rates for Proposed $PM_{2.5}$ Risk Analysis Locations.*

Health Effect	Los Angeles[1]	Detroit[2]	Seattle[3]
Hospital Admissions (per 100,000 general population/year)			
Pneumonia admissions (all ages): ICD codes 480-486	---	*	---
COPD and asthma admissions (all ages): ICD codes 490-496	---	*	*
COPD and asthma admissions (0-19): ICD codes 490-496	60	---	---
COPD and asthma admissions (20-64): ICD codes 490-496	113	---	---
COPD and asthma admissions (65 and older): ICD codes 490-496	158	---	---
Cardiovascular admissions (20-64): ICD codes: 390-429	428	---	---
Cardiovascular admissions (65 and over): ICD codes: 390-429	776	---	---
Ischemic heart disease (all ages): ICD codes 410-414	---	*	---
Dysrhythmia (all ages): ICD code 427	---	*	---
Congestive heart failure (all ages): ICD code 428	---	*	---

Hospitalization rates are presented only for the locations in which the C-R functions were estimated. See Exhibit C.2 in Appendix C for $PM_{2.5}$ C-R functions proposed to be used in the analyses. All incidence rates are rounded to the nearest unit.

* Hospitalization data for Seattle are being obtained and we are pursuing the availability of hospitalization data for Detroit.

1. Los Angeles County. The numbers of hospitalization discharges in 1999 were obtained from California's Office of Statewide Health Planning and Development for Los Angeles County. The number of discharges was divided by the 1999 population from U.S. Census estimates to obtain rates.

2. Wayne County.

3. King County.

In the absence of other sources of baseline incidence data for respiratory symptoms that do not require hospitalization, baseline rates for these health endpoints will be taken from the studies used to generate the C-R functions proposed for use in the risk analysis.

5.2. Sources of incidence data for the $PM_{10-2.5}$ risk analyses [to be added, if appropriate]

6. Sources of Uncertainty

The PM health risk models that will be used in the risk analyses will combine information about PM for specific urban areas to derive estimates of the annual incidence of specified health effects associated with "as is" PM concentrations and the reduction in incidence that would result upon just meeting specified PM standards in those areas. The three main inputs to such analyses -- air quality information, C-R information, and baseline incidence and population information -- all vary from one time and location to another time and location. It is rarely possible to obtain complete information that is specific to the assessment periods and proposed analysis locations on *all* of these input components.

For some components of the analyses (e.g., air quality information) it is possible to obtain location-specific information for many, but not all days in the year. Some uncertainty surrounding the results of the analyses will therefore arise from the incompleteness of such data. Even if air quality data were complete, since PM concentrations are measured, there is always some degree of measurement error.

For other components of the analyses (e.g., baseline incidence rates), it may not be possible to obtain any information that is specific to the analysis periods and locations. For these components, it will be necessary to rely on information from other times and/or locations. This will result in additional uncertainty surrounding the results of the analyses.

Finally, even if the input values are from the same times and locations as the analysis periods and locations, they will be only *estimates*, and will therefore have statistical uncertainty, including sampling error, surrounding them. The specific sources of uncertainty in the proposed PM risk analyses are described in detail below and are summarized in Exhibit 6.1.

Although the PM risk analyses will consider both mortality and a variety of morbidity health effects, not all health effects which may result from PM exposure will be included. Only those for which there was sufficient epidemiological evidence from studies which met the study selection criteria (see Section 3) are proposed to be included in the risk analyses. Other possible health effects reported to be associated with short- and/or long-term exposures to $PM_{2.5}$ and $PM_{10-2.5}$ will be considered qualitatively in the OAQPS Staff Paper. Thus, the proposed risk analyses will not represent all of the health risks associated with PM exposures.

For respiratory symptoms and illnesses, the number of cases avoided will be based only on the age group under study. For example, lower respiratory symptoms were examined in Schwartz and Neas (2000) for children ages 7-14. It is likely that the effect of PM on lower respiratory symptoms does not begin at age 7 and end at age 14; however, data are not available to estimate the number of cases avoided for other age groups. Therefore, a substantial number of potentially avoided health effects are omitted from this analysis.

Exhibit 6.1. Key Uncertainties in the Risk Analyses

Uncertainty	Direction of Potential Error	Comments
Empirically estimated C-R relations	?	Statistical association does not prove causation. Because C-R functions are empirically estimated, there is uncertainty surrounding these estimates. Omitted confounding variables could cause bias in the estimated PM coefficients.
Functional form of C-R relation	?	Statistical significance of coefficients in an estimated C-R function does not necessarily mean that the mathematical form of the function is the best model of the true C-R relation.
Lag structure of C-R relation	-	There is some evidence of a distributed lag. Most models, however, included only one lag. Omitted lags could cause downward bias in the predicted incidence associated with a given reduction in PM concentrations. A comparison or sensitivity analysis using an approach to estimate the possible impact of using a distributed lag C-R function is proposed.
Transferability of C-R relations	?	C-R functions may not provide an adequate representation of the C-R relation in times and places other than those in which they were estimated. For example, populations in the analysis locations may have more or fewer members of sensitive subgroups than locations in which functions were derived, which would introduce additional uncertainty related to the use of a given C-R function in the analysis location. However, in the majority of cases, the proposed risk analyses will rely on C-R functions estimated from studies conducted in the same location.

Uncertainty	Direction of Potential Error	Comments
Extrapolation of C-R relations beyond the range of observed PM data	?	A C-R relation estimated by an epidemiological study may not be valid at concentrations outside the range of concentrations observed during the study. To partially address this problem, risk will not be calculated for PM levels below the lowest observed level in a study, if it's reported. However, not all studies report the range of PM concentrations observed. If the lowest observed level is not reported, risk will be estimated down to background level, which may be lower than the lowest PM level observed in the study.
Truncation of risk estimates at the lowest PM concentration observed in a study	-	To avoid relying on a C-R function below the lowest PM concentration from which it was estimated, risk will not be calculated for PM levels below the lowest observed level in a study, if it's reported. If there is any positive relation between PM and the health response below this level, this procedure will understate the PM impact.
Adequacy of PM characterization	?	Only size differentiated particle mass per unit volume has been explicitly considered, and not, for example, chemical composition. However, in the majority of cases, the proposed risk analyses will rely on C-R functions estimated from studies conducted in the same location and, therefore, capture to some extent any potential impact on health effects due to differences in composition.
Accuracy of PM mass measurement	?	Possible differences in measurement error, losses of particular components, and measurement method between the assessment locations and the study locations would be expected to add uncertainty to quantitative estimates of risk.
Adjustment of air quality distributions to simulate just meeting alternative standards	?	The pattern and extent of daily reductions in PM concentrations that would result if alternative PM standards were just met is not known. Although the assumption that PM concentrations would be reduced by the same percentage on all days appears reasonable given the patterns observed based on historical data, there remains uncertainty about the shape of the air quality distribution of daily levels upon just meeting alternative PM standards which will depend on future air quality control strategies.

Uncertainty	Direction of Potential Error	Comments
Background PM concentrations	?	The calculation of PM risk associated with "as is" air quality and of risk reductions that would result if alternative standards were just met requires as inputs the background PM concentrations in each of the assessment locations. Background concentrations were estimated for the eastern and western regions of the country, but not specifically for the assessment locations. In addition, a constant value is proposed to be used for the estimated background, which will not take into account seasonal or daily variability in background concentrations. Therefore, there is uncertainty associated with the estimated background concentrations that will be used.
Baseline health effects data	?	Data on baseline incidence is uncertain for a variety of reasons. For example, location- and age-group-specific baseline rates may not be available in all cases. Baseline incidence may change over time for reasons unrelated to PM.

6.1. Concentration-response functions

The C-R function is a key element of the PM risk analyses. The quality of the risk analyses depends, in part, on (1) how well the C-R functions used in the risk analyses have been estimated (e.g., whether they are unbiased estimates of the relation between the population health response and ambient PM concentration in the study locations), (2) how applicable these functions are to the analysis periods and locations, and (3) the extent to which these relations apply beyond the range of the PM concentrations from which they were estimated. These issues are discussed in the subsections below.

6.1.1. Uncertainty associated with the estimated concentration-response functions in the study locations

The uncertainty associated with an estimate of a C-R function reported by a study depends on the sample size and the study design. The 2001 draft PM CD has evaluated the substantial body of PM epidemiological studies. In general, critical considerations in evaluating the design of an epidemiological study include the adequacy of the measurement of average ambient PM, the adequacy of the health effects incidence data, and the consideration of potentially important health determinants and causal (confounding) factors such as:

- other pollutants;
- exposure to other health risks, such as smoking and occupational exposure; and
- demographic characteristics, including age, sex, socioeconomic status, and access to medical care.

The list of proposed studies for inclusion in the PM risk analyses has been guided by the evaluations in the 2001 draft PM CD. Two of the criteria for selecting studies to be used in the PM risk analyses address the adequacy of the measurement of average ambient PM. One criterion is that $PM_{2.5}$ was measured rather than estimated on a reasonable proportion of the days in the study. Another criterion is that the measure of PM used in the study was $PM_{2.5}$ or $PM_{2.1}$. These two criteria are designed to minimize measurement error in the estimated PM coefficients in the C-R functions used in the risk analyses.

To the extent that a study did not address all critical factors, there is uncertainty associated with the C-R function estimated in that study. It may result in either over- or underestimates of risk associated with ambient PM concentrations in the location in which the study was carried out. Techniques for addressing the problem of confounding factors and other study design issues have improved over the years, however, and the epidemiological studies currently available for use in the PM risk analyses provide a higher level of confidence in study quality than ever before.

When a study is conducted in a single location, the problem of possible confounding co-pollutants may be particularly difficult, if co-pollutants are highly correlated in the study location. Single-pollutant models, which omit co-pollutants, may produce overestimates of the PM effect, if some of the effects of other pollutants (omitted from the model) are falsely attributed to PM. With regard to gaseous co-pollutants as potential confounders, a new multi-city study (NMMAPS; Samet et al., 2000) has evaluated the effects of PM_{10} alone and in combination with each of the monitored gaseous co-pollutants across the 90 largest U.S. cities and reported that associations found between PM_{10} and mortality were not confounded by the presence of the gaseous co-pollutants. (preliminary draft PM SP, p. 3-18) On the other hand, statistical estimates of a PM effect based on a multi-pollutant model can be more uncertain, and even statistically insignificant, if the co-pollutants included in the model are highly correlated with PM. This means that, although the expected value of the estimated PM coefficient is correct, the estimate based on any particular sample may be too low or too high. As a result of these considerations, we plan to report estimates based on multi-city studies that used $PM_{2.5}$ as the indicator, when available, as well as estimates based on the single-city study conducted in the risk analysis urban area.

6.1.2. Applicability of concentration-response functions in different locations

As described in Section 3, risk analysis locations have been selected on the basis of where C-R functions have been estimated, to avoid the uncertainties associated with applying a C-R function estimated in one location to another location. However, the PM risk analyses may also use a C-R function that was (1) estimated in several different locations (in a multi-city study) or, in some limited cases, may apply a C-R function from a different location as part of a sensitivity analysis. The accuracy of the results based on such multi-city or other location C-R functions rests in part on the "transferability" of the C-R relation from one location to another. That is, it rests on the assumption that the relation between ambient PM and a given population health response is the same in the two locations.

The relation between ambient PM concentration and the incidence of a given health endpoint in the population (the population health response), the C-R relation, depends on (1) the relation between ambient PM concentration and personal exposure to ambient-generated PM and (2) the relation between personal exposure to ambient-generated PM and the population health response. Both of these are likely to vary to some degree from one location to another.

The relation between ambient PM concentration and personal exposure to ambient-generated PM will depend on patterns of behavior, such as the amount of time spent outdoors, as well as on factors affecting the extent to which ambient-generated PM infiltrates into indoor environments. The relation between personal exposure to ambient-generated PM and the population health response will depend on both the composition of the PM and on the composition of the population exposed to it.

The composition of PM (e.g., the proportion that is fine particles versus coarse particles and the chemical constituents of the PM) is known to differ from one location to another. (For example, in some locations, PM is mostly fine particles; in other locations, it is mostly coarse windblown dust.) As discussed in the 2001 draft PM CD recent studies provide some evidence for health effect associations with many different PM components, including sulfates, acids, and metals. However, as stated in the preliminary draft PM SP (p.3-80), "the evidence is still too limited to allow identification of which PM components or sources might be more toxic than others, and growing evidence indicates that there are numerous potentially toxic PM components and there may also be interaction occurring between components."

Exposed populations also differ from one location to another in characteristics that are likely to affect their susceptibility to PM air pollution. For instance, people with pre-existing conditions such as chronic bronchitis are probably more susceptible to the adverse effects of exposure to PM, and populations vary from one location to another in the prevalence of specific diseases. Also, some age groups may be more susceptible than others, and population age distributions also vary from one location to another. Closely matching populations observed in studies to the populations of the assessment locations is not possible for many characteristics (for example, smoking status, workplace exposure, socioeconomic status, and the prevalence of highly susceptible subgroups).

Other pollutants, such as carbon monoxide and ozone, may also play a role in causing health effects, either independently or in combination with PM. Inter-locational differences in these pollutants could also induce differences in the C-R relation between one location and another.

In summary, the C-R relation is most likely not the same everywhere. Even if the relation between personal exposure to ambient-generated PM and population health response were the same everywhere, the relation between ambient concentrations and personal exposure to ambient-generated PM may differ among locations. Similarly, even if the relation between ambient concentrations and personal exposure to ambient-generated PM were the same everywhere, the relation between personal exposure to ambient-generated PM and population health response may differ among locations. In either case, the C-R relation would differ.

6.1.3. Extrapolation beyond observed air quality levels

Although a C-R function describes the theoretical relation between ambient PM and a given health endpoint for all possible PM levels (down to zero), the estimation of a C-R function is based on real ambient PM values that are limited to the range of PM concentrations in the location in which the study was conducted. Thus, uncertainty in the shape of the estimated C-R function increases considerably outside the range of PM concentrations observed in the study.

The planned risk analyses will assume that the estimated C-R functions adequately represent the true C-R relation down to background levels in the assessment locations, in cases in which this background level is above the lowest concentrations used to derive the C-R functions. Estimates of risk will not be generated for concentrations below the minimum concentrations observed in the studies. Although the current $PM_{2.5}$ standards, and any alternatives that EPA is likely to consider, generally lie in the middle range of pollution levels observed in epidemiological studies, applying proportional rollbacks to the concentration distributions in the assessment locations may result in some modeled PM concentrations below the lowest levels observed in the studies. In such cases, the change in PM will be taken to be the difference between the "as is" levels and the lowest observed level in the study. This procedure avoids relying on a C-R function below the level of PM concentrations from which it was estimated. However, it will tend to understate the impact of just meeting alternative standards if there is actually a C-R relation below these lowest observed PM levels.

It is possible that there is a minimum concentration (i.e., threshold) below which PM is not associated with health effects. If there is such a concentration, including incidence reductions associated with reducing PM levels below this minimum threshold level in the total incidence reduction would overstate the risk attributable to PM or the incidence reductions that would result from just attaining a standard. Sensitivity analyses will examine the sensitivity of the results of the risk analyses to different assumptions about potential thresholds.

The C-R relation may also be less certain towards the upper end of the concentration range being considered in a risk analysis, particularly if the PM concentrations in the assessment location exceed the PM concentrations observed in the study location. Even though it may be reasonable to model the C-R relation as log-linear over the ranges of PM concentrations typically observed in epidemiological studies, it may not be log-linear over the entire range of PM levels at the locations considered in the PM risk analyses.[22]

6.2. The air quality data

6.2.1. Use of $PM_{2.5}$ as the indicator

PM is measured in units of mass per unit volume, typically in micrograms per cubic meter. The PM risk analyses will use PM size classes -- e.g., $PM_{2.5}$, and if appropriate $PM_{10-2.5}$, and the chemical composition of PM will not be considered explicitly in any of the risk analyses (as it was not in most of the epidemiological studies used in these analyses). As summarized in Chapter 9 of the 2001 draft PM CD, recent studies provide new evidence for health effects associations with many different PM components. Recognizing that ambient PM exposure has

[22]Although the C-R functions are log-linear, they are practically linear. It is still unlikely, however, that a linear function is appropriate over a very wide range of PM concentrations.

been associated with increases in numerous health indices, the evidence is still too limited to allow identification of which PM components or sources might be more toxic than others, and growing evidence indicates that there are numerous potentially toxic PM components and some components may act in combination (preliminary PM SP, p.3-80). It is possible that PM risks may differ from one area to another with differing PM composition, but this potential source of uncertainty cannot be tested in these risk analyses. However, because the proposed risk analyses primarily will use C-R functions estimated from studies conducted in the same location as the analysis location, the C-R functions already capture to some extent the potential impact of differential composition. To the extent that composition differentially affects toxicity and if future control strategies alter the composition in an area, then this introduces an additional uncertainty into the risk estimates associated with just meeting the current or alternative PM standards.

6.2.2. Adequacy of PM air quality data

The method of averaging data from monitors across a metropolitan area in the risk analyses is similar to the methods used to characterize ambient air quality in most of the epidemiology studies. Ideally, the measurement of average daily ambient PM concentrations in the study location are unbiased. In this case, unbiased risk predictions in the assessment location depend, in part, on an unbiased measurement of average daily ambient PM concentrations in the assessment location as well. If, however, the measurement of average daily ambient PM concentrations in the study location are biased, unbiased risk predictions in the assessment location are still possible if the measurement of average daily ambient PM concentrations in the assessment location incorporate the same bias as exists in the study location measurements. Because this is not known, however, the adequacy of the PM measurements in the assessment locations is a source of uncertainty in the risk analysis.

PM air quality data are not available for all days of the year chosen for risk analysis in any of the assessment locations. The change in the incidence of a health effect over the course of the year corresponding to a given change in daily PM levels is calculated based on the assumption that PM levels on those days with PM data are representative of levels on those days without PM data (see Section 2.6 for an explanation of the method of extrapolating changes in health effects incidence to an entire year). If there are seasonal differences in average PM levels and in monitoring frequencies, a simple annual adjustment for missing data could result in a biased estimate of total annual incidence change. To minimize the presence of bias due to an uneven distribution of missing data throughout the year, incidence changes in different quarters of the year will be scaled separately, and the scaled quarterly results will be added.

Because the PM data in each assessment location are limited to a specific year, the results of the risk analyses will be generalizable to the present only to the extent that ambient PM levels in the available data are similar to current ambient PM levels in those locations. A substantial

difference between PM levels in the years used in the risk analyses and current PM levels could imply a substantial difference in predicted incidences of health effects. This is not expected to be a large problem for the $PM_{2.5}$ risk analyses, however, because adequate $PM_{2.5}$ monitoring data are available for each of the proposed assessment locations in quite recent years (2000 for all locations except for Phoenix, and 1999 for Phoenix).

6.2.3. Simulation of reductions in PM concentrations to just meet alternative standards

The pattern of daily PM concentrations that would result if alternative PM standards were just met in any of the assessment locations is, of course, not known. Although the assumption that PM concentrations will be reduced by the same percentage on all days may be a reasonable approximation, it is only an approximation. There is therefore uncertainty surrounding the predicted daily changes in PM concentrations that would result if alternative standards were just met, and consequently uncertainty surrounding the associated daily changes in population health response.

6.3. Baseline health effects incidence rates

Most of the C-R functions to be used in the PM risk analyses are log-linear (see equations 1 through 3 in Section 2.5). Given this functional form, the percent change in incidence of a health effect corresponding to a change in PM depends only on the *change* in PM levels (and not the actual value of either the initial or final PM concentration). This percent change is multiplied by a baseline incidence in order to determine the change in health effects incidence, as shown in equation 3 in Section 2.5:

$$\Delta y = y[e^{\beta \Delta x} - 1] . \qquad (3\text{-}1)$$

in which $e^{\beta \Delta x}$ is the relative risk, and [$e^{\beta \Delta x}$ - 1] is the percent change associated with a change in PM of Δx. If there has been an increase in PM (i.e., if Δx positive), then the relative risk will be greater than 1.0. If, for example, the relative risk associated with a change in PM of Δx is 1.05, then the percent change in incidence of the health effect is 0.05 (5%). The change in incidence of the health effect associated with a change in PM of Δx is, then, 5 percent of the baseline incidence, y. Predicted changes in incidence therefore depend on the baseline incidence of the health effect.

6.3.1. Quality of incidence data

County-specific incidence data are available for mortality for all counties. We are currently in the process of obtaining hospital admissions baseline incidence data for Seattle and investigating the availability of such data for Detroit and Boston. This is clearly preferable to using non-local data, such as national incidence rates. As with any health statistics, however, misclassification of disease, errors in coding, and difficulties in correctly assigning residence location are potential problems. These same potential sources of error are present in most epidemiological studies. In most cases, the reporting institutions and agencies utilize standard forms and codes for reporting, and quality control is monitored.

Data on hospital admissions are actually hospital discharge data rather than admissions data. Because of this, the date associated with a given hospital stay is the date of discharge rather than the date of admissions. Therefore, there may be some hospital admissions in an assessment location in the year of interest (e.g., 1999) that are not included in the baseline incidence rate, if the date of discharge was after the year ended, even though the date of admissions was within the year. Similarly, there may be some hospital admissions that preceded the year of interest that are included in the baseline incidence rate because the date of discharge was within the year of interest. This is a very minor problem, however, partly because the percentage of such cases is likely to be very small, and partly because the error at the beginning of the year (i.e., admissions that should not have been included but were) will largely cancel the error at the end of the year (i.e., admissions that should have been included but were not).

Another minor uncertainty surrounding the hospital admissions baseline incidence rates arises from the fact that these rates are based on the reporting of hospitals within each of the assessment counties. Hospitals report the numbers of ICD code-specific discharges in a given year. If people from outside the county use these hospitals, and/or if residents of the county use hospitals outside the county, these rates will not accurately reflect the numbers of county residents who were admitted to the hospital for specific illnesses during the year, the rates that are required for the risk analyses. Once again, however, this is likely to be a very minor problem because the health conditions studied tend to be acute events that require immediate hospitalization, rather than planned hospital stays.

When local incidence data are not available, national rates will be used if possible. Estimates of national rates are generally considered reliable, due to the large sample sizes on which they are based. As the source population becomes smaller and the event rarer, the reliability may decrease, due the infrequency of occurrence. Most endpoints considered in these analyses, however, are common occurrences, and national sample sizes should be substantial. There is still uncertainty, however, about the extent to which a national rate is an adequate representation of a local rate.

Incidence rates for respiratory symptoms will be obtained from the study reporting the C-R function for those endpoints, Schwartz and Neas (2000). The baseline incidence rates reported in that study were based on all locations combined. Therefore there is some uncertainty associated with applying it to the individual locations. In addition, because this study is a reanalysis of data collected earlier, changes in baseline incidence rates over time could introduce additional uncertainty into the analysis.

Regardless of the data source, if actual incidence rates are higher than the incidence rates used, risks will be underestimated. If incidence rates are lower than the incidence rates used, then risks will be overestimated. For most of the C-R functions, the incidence rates affect the estimation of the changes in the number of cases associated with changes in PM (see equation 3 in Section 2.5), but not the estimation of the percentage changes in PM-related cases. The uncertainties in identifying the correct baseline incidence rates therefore affect only one portion of the results.

Both morbidity and mortality rates change over time for various reasons. One of the most important of these is that population age distributions change over time. The old and the extremely young are more susceptible to many health problems than is the population as a whole. The most recent available data will be used in the risk analyses. However, the average age of the population in many locations will increase as post-WWII children age. Consequently, the baseline incidence rates for some endpoints may rise, resulting in an increase in the number of cases attributable to any given level of PM pollution. Alternatively, areas which experience rapid in-migration, as is currently occurring in the South and West, may tend to have a decreasing mean population age and corresponding changes in incidence rates and risk. Temporal changes in incidence are relevant to both morbidity and mortality endpoints. However, the most recent available data will be used in all cases, so temporal changes are not expected to be a large source of uncertainty.

6.3.2. Lack of daily health effects incidence rates

Both ambient PM levels and the daily health effects incidence rates corresponding to ambient PM levels vary somewhat from day to day. Those risk analyses based on C-R functions estimated by short-term exposure studies calculate daily changes in incidence and sum them over the days of the year to predict an annual change in health effect incidence. However, only annual baseline incidence rates are available. Average daily baseline incidence rates, necessary for short-term daily C-R functions, will be calculated by dividing the annual rate by 365. To the extent that PM affects health, however, actual incidence rates would be expected to be somewhat higher than average on days with high PM concentrations; using an average daily incidence rate would therefore result in underestimating the changes in incidence on such days. Similarly, actual incidence rates would be expected to be somewhat lower than average on days with low PM concentrations; using an average daily incidence rate would therefore result in

overestimating the changes in incidence on low PM days. Both effects would be expected to be small, however, and should largely cancel one another out.

References

Abt Associates Inc. July 3, 1996 (Revised November 1996). "A Particulate Matter Risk Assessment for Philadelphia and Los Angeles." Prepared for the Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency, Contract No. 68-W4-0029. Available electronically on the web at: www.epa.gov/ttn/oarpg/t1sp.html.

Abt Associates Inc. June 5, 1997a. Revision of Mortality Incidence Estimates Based on Pope et al. (1995) in the Abt Particulate Matter Risk Assessment Report. Memorandum from Ellen Post and John Voyzey, Abt Associates Inc. to John Bachmann, Allyson Siwik, Michele McKeever, and Harvey Richmond, U.S. EPA/OAQPS.

Abt Associates Inc. June 6, 1997b. Revision of Mortality Incidence Estimates Based on Pope et al. (1995) in the December 1996 Supplement to the Abt Particulate Matter Risk Assessment Report. Memorandum from Ellen Post, Abt Associates Inc. to John Bachmann, Allyson Siwik, Michele McKeever, and Harvey Richmond, U.S. EPA/OAQPS.

Burnett, R.T., S. Cakmak, J.R. Brook and D. Krewski. 1997. The role of particulate size and chemistry in the association between summertime ambient air pollution and hospitalization for cardiorespiratory diseases. *Environ Health Perspect.* Vol. 105(6): 614-20.

Burnett, R.T., S. Cakmak, M.E. Raizenne, D. Stieb, R. Vincent, D. Krewski, J.R. Brook, O. Philips and H. Ozkaynak. 1998. The association between ambient carbon monoxide levels and daily mortality in Toronto Canada. Journal of the Air & Waste Management Association. Vol. 48(8): 689-700.

Burnett, R.T., M. Smith-Doiron, D. Stieb, S. Cakmak and J.R. Brook. 1999. Effects of particulate and gaseous air pollution on cardiorespiratory hospitalizations. Archives Environmental Health. Vol. 54(2): 130-139.

Burnett, R.T., J. Brook, T. Dann, C. Delocla, O. Philips, S. Cakmak, R. Vincent, M.S. Goldberg and D. Krewski. 2000. Association between Particulate and Gas-Phase Components of Urban Air Pollution and Daily Mortality in Eight Canadian Cities. *Inhalation Toxicology*. Vol. 12(Supplement 4): 15-39.

Deck, L. B., E. S. Post, E. Smith, M. Wiener, K. Cunningham, and H. Richmond. 2001 Estimates of the Health Risk Reductions Associated with Attainment of Alternative Particulate Matter Standards in Two U.S. Cities. *Risk Analysis* Vol. 21(5): 821-835.

Delfino, R.J., A.M. Murphy-Moulton, R.T. Burnett, J.R. Brook and M.R. Becklake. 1997. Effects of air pollution on emergency room visits for respiratory illnesses in Montreal, Quebec. *Am J Respir Crit Care Med.* Vol. 155(2): 568-576.

Dockery, D.W., F.E. Speizer, D.O. Stram, J.H. Ware, J.D. Spengler and B.G. Ferris, Jr. 1989. Effects of Inhalable Particles on Respiratory Health of Children. *Am Rev Respir Dis.* Vol. 139: 587-594.

Dockery, D.W., J. Schwartz and J.D. Spengler. 1992. Air Pollution and Daily Mortality - Associations With Particulates and Acid Aerosols. *Environmental Research.* Vol. 59(2): 362-373.

Dockery, D.W., J. Cunningham, A.I. Damokosh, L.M. Neas, J.D. Spengler, P. Koutrakis, J.H. Ware, M. Raizenne and F.E. Speizer. 1996. Health Effects of Acid Aerosols On North American Children - Respiratory Symptoms. *Environ Health Perspect.* Vol. 104(5): 500-505.

Fairley, D. 1999. Daily mortality and air pollution in Santa Clara County, California: 1989-1996. *Environ Health Perspect.* Vol. 107(8): 637-41.

Goldberg, M.S., J.C. Bailar III, R.T. Burnett, J.R. Brook, R. Tamblyn, Y. Bonvalot, P. Ernst, K.M. Flegel, R.K. Singh, M. and Valois. Identifying Subgroups of the General Population That May Be Susceptible to Short-Term Increases in Particulate Air Pollution: A Time-Series Study in Montreal, Quebec. Health Effects Institute Research Report, Number 97, October 2000.

Ito, K., P. L. Kinney, and G.D. Thurston 1995. Variations in PM-10 Concentrations Within Two Metropolitan Areas and Their Implications For Health Effects Analyses. *Inhalation Toxicology* 7(5): 735-745.

Laden, F., L.M. Neas, D.W. Dockery and J. Schwartz. 2000. Association of fine particulate matter from different sources with daily mortality in six U.S. cities. *Environ Health Perspect.* Vol. 108(10): 941-7.

Lipfert, F.W., S.C. Morris and R.E. Wyzga. 2000. Daily mortality in the Philadelphia metropolitan area and size-classified particulate matter. *J Air Waste Manag Assoc.* Vol. 50(8): 1501-13.

Lippmann, M., K. Ito, A. Náádas, and R.T. Burnett. 2000. Association of Particulate Matter Components with Daily Mortality and Morbidity in Urban Populations. Health Effects Institute Research Report, Number 95.

Mar, T.F., G.A. Norris, J.Q. Koenig and T.V. Larson. 2000. Associations between air pollution and mortality in Phoenix, 1995-1997. *Environ Health Perspect*. Vol. 108(4): 347-53.

McConnell, R., K. Berhane, F. Gilliland, S.J. London, H. Vora, E. Avol, W.J. Gauderman, H.G. Margolis, F. Lurmann, D.C. Thomas and J.M. Peters. 1999. Air pollution and bronchitic symptoms in Southern California children with asthma. *Environ Health Perspect*. Vol. 107(9): 757-60.

Moolgavkar, S.H. 2000a. Air Pollution and Daily Mortality in Three U.S. Counties. *Environ Health Perspect*. Vol. 108(8): 777-784.

Moolgavkar, S.H. 2000b. Air pollution and hospital admissions for diseases of the circulatory system in three U.S. metropolitan areas. *J Air Waste Manag Assoc*. Vol. 50(7): 1199-206.

Moolgavkar, S.H. 2000c. Air Pollution and Hospital Admissions for Chronic Obstructive Pulmonary Disease in Three Metropolitan Areas in the United States. *Inhalation Toxicology*. Vol. 12(Supplement 4): 75-90.

Moolgavkar, S.H., W. Hazelton and G. Luebeck. 2000. Air Pollution, Pollens, and Admissions for Chronic Respiratory Disease in King County, Washington. *Inhalation Toxicology*. Vol. 12(Supplement 1): 157-171.

Norris, G., S.N. YoungPong, J.Q. Koenig, T.V. Larson, L. Sheppard and J.W. Stout. 1999. An association between fine particles and asthma emergency department visits for children in Seattle. Environ Health Perspect. Vol. 107(6): 489-93.

Ostro, B. 1995. Fine particulate air pollution and mortality in two Southern California counties. Environmental Research. Vol. 70(2): 98-104.

Ostro, B.D., R. Broadwin and M.J. Lipsett. 2000. Coarse and fine particles and daily mortality in the Coachella Valley, California: a follow-up study. J Expo Anal Environ Epidemiol. Vol. 10(5): 412-9.

Peters, J.M., E. Avol, W. Navidi, S.J. London, W.J. Gauderman, F. Lurmann, W.S. Linn, H. Margolis, E. Rappaport, H. Gong and D.C. Thomas. 1999. A Study of Twelve Southern California Communities with Differing Levels and Types of Air Pollution. I. prevalence of respiratory morbidity. *Am J Respir Crit Care Med*. Vol. 159(3): 760-767.

Peters, A., D.W. Dockery, J.E. Muller and M.A. Mittleman. 2001. Increased particulate air pollution and the triggering of myocardial infarction. Circulation. Vol. 103(23): 2810-5.

Post, E., L. Deck, K. Larntz, D. Hoaglin. 2001. An Application of an Empirical Bayes Estimation Technique to the Estimation of Mortality Related to Short-Term Exposure to Particulate Matter. *Risk Analysis Vol. 21(5):* 837-842.

Romieu, I., F. Meneses, S. Ruiz, J.J. Sienra, J. Huerta, M.C. White and R.A. Etzel. 1996. Effects of air pollution on the respiratory health of asthmatic children living in Mexico City. Am J Respir Crit Care Med. Vol. 154(2 Pt 1): 300-7.

Samet, J.M., S.L. Zeger, F. Dominici, F. Curriero, I. Coursac, D.W. Dockery, J. Schwartz, and A. Zanobetti. 2000. The National Morbidity, Mortality, and Air Pollution Study, Part II: Morbidity, Mortality, and Air Pollution in the United States. Health Effects Institute Research Report, Number 94, Part II.

Schwartz, J. 2000. Harvesting and long-term exposure effects in the relationship between air pollution and mortality. *Am. J. Epidemiology* 151: 440-448.

Schwartz, J., D.W. Dockery, L.M. Neas, D. Wypij, J.H. Ware, J.D. Spengler, P. Koutrakis, F.E. Speizer and B.G. Ferris. 1994. Acute Effects of Summer Air Pollution On Respiratory Symptom Reporting in Children. *Am J Respir Crit Care Med*. Vol. 150(5): 1234-1242.

Schwartz, J., D.W. Dockery and L.M. Neas. 1996. Is Daily Mortality Associated Specifically With Fine Particles. *Journal of the Air & Waste Manag Assoc*. Vol. 46(10): 927-939.

Schwartz, J. and L.M. Neas. 2000. Fine particles are more strongly associated than coarse particles with acute respiratory health effects in schoolchildren [see comments]. *Epidemiology*. Vol. 11(1): 6-10.

Smith, R.L., D. Spitzner, Y. Kim and M. Fuentes. 2000. Threshold dependence of mortality effects for fine and coarse particles in Phoenix, Arizona. *J Air Waste Manag Assoc*. Vol. 50(8): 1367-79.

Stieb, D.M., R.C. Beveridge, J.R. Brook, M. Smith-Doiron, R.T. Burnett, R.E. Dales, S. Beaulieu, S. Judek and A. Mamedov. 2000. Air pollution, aeroallergens and cardiorespiratory emergency department visits in Saint John, Canada. *J Expo Anal Environ Epidemiol*. Vol. 10(5): 461-77.

Thurston, G.D., K. Ito, C.G. Hayes, D.V. Bates and M. Lippmann. 1994. Respiratory hospital admissions and summertime haze air pollution in Toronto, Ontario: consideration of the role of acid aerosols. *Environ Res*. Vol. 65(2): 271-290.

Tsai, F.C., M.G. Apte and J.M. Daisey. 2000. An Exploratory Analysis of the Relationship between Mortality and the Chemical Composition of Airborne Particulate Matter. *Inhalation Toxicology*. Vol. 12(Supplement 2): 121-135.

U.S. Environmental Protection Agency (U.S. EPA). 1996a. *Air Quality Criteria for Particulate Matter*, (EPA 600/P-95/001aF), 3v, National Center for Environmental Assessment, Office of Research and Development, Research Triangle Park, NC.

U.S. Environmental Protection Agency (U.S. EPA) July 1996b. *Review of the National Ambient Air Quality Standards for Particulate Matter: Policy Assessment of Scientific and Technical Information - OAQPS Staff Paper,* (EPA/452/R-96-013), Office of Air Quality Planning and Standards, Research Triangle Park, NC 27711. Available from: NTIS, Springfield, VA; PB97-115406REB or electronically on the internet at: http://www.epa.gov/ttn/oarpg/t1sp.html.

U.S. Environmental Protection Agency (U.S. EPA) March 2001a. *Air Quality Criteria for Particulate Matter*, Second External Review Draft (EPA 600/P-99/002aB), 2v, National Center for Environmental Assessment, Office of Research and Development, Research Triangle Park, NC. Available electronically on the internet at: http://www.epa.gov/ncea/partmatt.htm.

U.S. Environmental Protection Agency (U.S. EPA) June 2001b. *Review of the National Ambient Air Quality Standards for Particulate Matter: Policy Assessment of Scientific and Technical Information - OAQPS Staff Paper,* Preliminary Draft, Office of Air Quality Planning and Standards, Research Triangle Park, NC. Available electronically on the internet at: http://www.epa.gov/ttn/oarpg/t1sp.html.

Wolff, G.T.. 1996. Letter from George T. Wolff, Chair, Clean Air Scientific Advisory Committee. Closure by the Clean Air Scientific Advisory Committee (CASAC) on the Staff Paper for Particulate Matter. June 13.

Appendix A. Air Quality Assessment: The PM Data

This Appendix describes the $PM_{2.5}$ data for the urban counties proposed for use in the risk analyses (see Section 3 for selection of locations). The average ambient $PM_{2.5}$ concentration in an assessment location on a given day is represented by the *average* of 24-hour average $PM_{2.5}$ levels at the different monitors in that location that reported on that day. This approach is consistent with what has been done in epidemiological studies estimating PM C-R functions. Also, because people are often quite mobile (e.g., living in one part of a county and working in another), an area-wide average PM level may be a more meaningful measure of ambient PM concentration than PM levels at individual monitors. Ito et al. (1995), for example, found that averaging PM_{10} concentrations reported at monitors in different places generally improved the significance of the association between PM_{10} and mortality in Chicago, compared with using individual monitors. If $PM_{10-2.5}$ risk analyses are also carried out, the same approach will be used. $PM_{10-2.5}$ will be calculated by subtracting $PM_{2.5}$ from PM_{10} at co-located monitors.

A.1. The $PM_{2.5}$ data

$PM_{2.5}$ data for each of the urban areas identified in Section 3 (Philadelphia, Philadelphia Region, Los Angeles, Phoenix, San Jose, Boston, Detroit, St. Louis, and Seattle) were obtained from EPA's Aerometric Information Retrieval System (AIRS) for the year 2000 (or 1999 where there was not adequate data for an area in the year 2000). In order for an urban area to be included in the risk analysis, the location must contain at least one monitor with 11 or more observations per quarter and 122 observations per year (1 in 3 day monitoring). Once the criteria for inclusion are met, all monitors with at least 11 observations per quarter will be used for each location. The cutoff of 11 observations per quarter is based on EPA guidance on measuring attainment of the daily and annual particulate matter standards outlined in Appendix N of the July 18, 1997 Federal Register Notice (available on the web at www.epa.gov/ttn/oarpg/t1pfpr.html). The guidance requires that at least 75 percent of the scheduled sampling days for each quarter have valid data. Based on a one in six day sampling protocol, the minimum required number of observations would be 11 per quarter.

The numbers of days of observations by monitor and at the composite monitor, by quarter and for the year, along with annual averages and 98th percentile concentrations, are given in Exhibits A.1 through A.9 for each of the proposed locations. The locations of the monitors in each urban area are mapped in Exhibits A.10 through A.18. In these exhibits the first five digits, which denote the FIPS code designation, are omitted in the legends. An initial check revealed insufficient $PM_{2.5}$ data for Phoenix in 2000, but substantially more data were available for 1999. Therefore, 1999 $PM_{2.5}$ data were collected for Phoenix for use in the analysis and 2000 $PM_{2.5}$ data for the remaining locations. The annual average at each monitor, and at the composite monitor, is the average of the four quarterly averages at the monitor. The 98th percentile at each monitor, and at the composite monitor, is calculated using the method used by EPA, as described in

Appendix N of the July 18, 1997 Federal Register Notice (available on the web at www.epa.gov/ttn/oarpg/t1pfpr.html). The only difference between the method proposed in the risk analyses and the standard EPA convention in calculating annual averages and 98th percentile values is that the EPA convention uses three years of data whereas the risk analyses will be based on only a single year of data (which is equivalent to assuming three identical years).

The maximum average of monitor-specific annual averages is used to determine the percent rollback necessary to meet an annual standard; the highest monitor-specific 98th percentile value is used to determine the percent rollback necessary to meet a daily standard. Although the composite monitor is not used in determining the percent rollback in the PM risk analyses, the percent rollback to simulate just meeting alternative standards is applied to the composite monitor.

Exhibit A.1. Number of Days on which $PM_{2.5}$ Concentration Data are Available, by Monitor and by Quarter, and $PM_{2.5}$ Concentrations. Los Angeles, 2000*

Monitor	Q1	Q2	Q3	Q4	Year Total	Annual Avg.	98th Percentile
AIRS 060370002881011	79	88	85	74	326	20.2	61.6
AIRS 060371103881011	59	66	67	83	275	21.5	73.9
AIRS 060371201881011	28	27	25	27	107	17.8	50.0
AIRS 060371301881011	28	29	30	29	116	23.2	62.8
AIRS 060371601881011	27	28	28	29	112	23.9	70.8
AIRS 060372005881011	30	28	27	25	110	19.4	54.0
AIRS 060374002881011	70	66	67	56	259	19.3	64.3
Composite[1]	85	90	92	90	357	20.8	68.9

*All concentrations are in $\mu g/m^3$; includes the section of Los Angeles County in the South Coast Air Basin. This excluded a single AIRS monitor in the Mojave Desert Air Basin.

1. The number of days at the composite monitor is the number of days on which at least one of the monitors reported.

Exhibit A.2. Number of Days on which $PM_{2.5}$ Concentration Data are Available, by Monitor and by Quarter, and PM2.5 Concentrations. Philadelphia, 2000*

Monitor	Q1	Q2	Q3	Q4	Year Total	Annual Avg.	98th Percentile
AIRS 421010004881011	61	45	76	84	266	14.7	37.6
AIRS 421010024881011	27	12	31	25	95	14.7	37.5
AIRS 421010136881011	54	45	72	75	246	14.4	41.5
Composite[1]	78	51	87	90	306	14.5	37.6

*All concentrations are in $\mu g/m^3$; includes Philadelphia County.

1. The number of days at the composite monitor is the number of days on which at least one of the monitors reported.

Exhibit A.3. Number of Days on which $PM_{2.5}$ Concentration Data are Available, by Monitor and by Quarter, and $PM_{2.5}$ Concentrations. Philadelphia Region, 2000*

Monitor	Q1	Q2	Q3	Q4	Year Total	Annual Avg.	98th Percentile
AIRS 340070003881011	29	28	29	22	108	15.0	32.1
AIRS 340071007881011	29	26	26	22	103	15.5	35.7
AIRS 340155001881011	25	28	30	25	108	15.1	34.1
AIRS 420170012881011	19	27	24	24	94	13.8	38.4
AIRS 420450002881011	27	28	30	27	112	16.0	36.2
AIRS 420910013881011	20	24	27	29	100	13.7	37.5
AIRS 421010004881011	61	45	76	84	266	14.7	37.6
AIRS 421010024881011	27	12	31	25	95	14.7	37.5
AIRS 421010136881011	54	45	72	75	246	14.4	41.5
Composite[1]	81	65	87	92	325	14.6	38.6

*All concentrations are in $\mu g/m^3$; includes Philadelphia, Bucks, Delaware, Montgomery, Camden (NJ), Gloucester (NJ), and Burlington Counties (NJ).

1. The number of days at the composite monitor is the number of days on which at least one of the monitors reported.

Exhibit A.4. Number of Days on which $PM_{2.5}$ Concentration Data are Available, by Monitor and by Quarter, and PM2.5 Concentrations. Phoenix, 1999*

Monitor	Q1	Q2	Q3	Q4	Year Total	Annual Avg.	98^{th} Percentile
AIRS 040139990881011	25	21	27	28	101	10.8	22.3
AIRS 040139991881011	43	69	87	64	263	13.1	32.1
AIRS 040139992881011	13	52	26	13	104	12.3	31.5
AIRS 040139997881011	28	74	91	65	258	11.7	26.1
Composite[1]	66	87	92	88	333	12.2	31.0

*All concentrations are in $\mu g/m^3$; includes Maricopa County.

Exhibit A.5. Number of Days on which $PM_{2.5}$ Concentration Data are Available, by Monitor and by Quarter, and $PM_{2.5}$ Concentrations. San Jose, 2000*

Monitor	Q1	Q2	Q3	Q4	Year Total	Annual Avg.	98^{th} Percentile
AIRS 060850004881012	76	14	15	74	179	13.5	56.6

*All concentrations are in $\mu g/m^3$; includes Santa Clara County.

Exhibit A.6. Number of Days on which $PM_{2.5}$ Concentration Data are Available, by Monitor and by Quarter, and $PM_{2.5}$ Concentrations. Boston, 2000*

Monitor	Q1	Q2	Q3	Q4	Year Total	Annual Avg.	98th Percentile
AIRS 250171102881011	21	22	13	11	67	8.9	26.8
AIRS 250250042881011	56	62	74	58	250	13.1	31.9
AIRS 250250043881011	31	26	17	13	87	15.8	35.2
Composite[1]	71	73	77	68	289	13.0	29.7

*All concentrations are in $\mu g/m^3$; includes Middlesex, Norfolk, and Suffolk Counties.

1. The number of days at the composite monitor is the number of days on which at least one of the monitors reported.

Exhibit A.7. Number of Days on which $PM_{2.5}$ Concentration Data are Available, by Monitor and by Quarter, and PM2.5 Concentrations. Detroit, 2000*

Monitor	Q1	Q2	Q3	Q4	Year Total	Annual Avg.	98^{th} Percentile
AIRS 261630001881011	81	86	87	85	339	15.6	38.6
AIRS 261630015881011	30	28	31	30	119	18.1	44.5
AIRS 261630016881011	83	74	78	89	324	15.4	40.1
AIRS 261630025881011	29	27	30	25	111	14.1	30.5
AIRS 261630033881011	28	23	27	29	107	19.9	43.3
AIRS 261630036881011	16	28	29	29	102	17.4	42.0
Composite[1]	90	89	92	92	363	16.0	37.0

*All concentrations are in $\mu g/m^3$; includes Wayne County.

1. The number of days at the composite monitor is the number of days on which at least one of the monitors reported.

Exhibit A.8. Number of Days on which $PM_{2.5}$ Concentration Data are Available, by Monitor and by Quarter, and $PM_{2.5}$ Concentrations. St. Louis, 2000*

Monitor	Q1	Q2	Q3	Q4	Year Total	Annual Avg.	98th Percentile
AIRS 171192009881011	31	30	30	29	120	16.0	36.3
AIRS 171634001881011	24	28	24	30	106	15.0	32.8
AIRS 290990012881011	31	24	18	30	103	14.8	27.4
AIRS 291831002881011	30	27	28	28	113	14.9	34.4
AIRS 291892003881011	29	30	31	28	118	14.8	30.8
AIRS 291895001881011	28	30	30	30	118	14.4	33.3
AIRS 295100085881011	88	89	92	89	358	16.4	34.8
AIRS 295100086881011	67	84	82	89	322	15.0	33.2
Composite[1]	91	91	92	92	366	15.7	34.1

*All concentrations are in $\mu g/m^3$; includes St. Louis, Franklin, Jefferson, St. Charles, Clinton (IL), Madison (IL), Monroe (IL), and St. Clair (IL) Counties and St. Louis City.

1. The number of days at the composite monitor is the number of days on which at least one of the monitors reported.

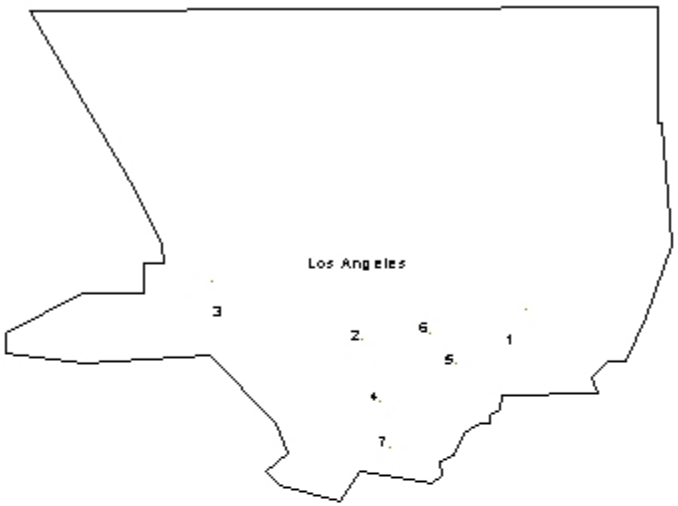
Exhibit A.9. Number of Days on which $PM_{2.5}$ Concentration Data are Available, by Monitor and by Quarter, and PM2.5 Concentrations. Seattle, 2000*

Monitor	Q1	Q2	Q3	Q4	Year Total	Annual Avg.	98th Percentile
AIRS 530330017881011	27	27	30	24	108	5.7	14.4
AIRS 530330021881011	82	91	88	90	351	11.9	35.1
AIRS 530330024881011	29	30	31	25	115	12.8	32.5
AIRS 530330027881011	28	28	29	29	114	9.4	23.9
AIRS 530330033881011	20	30	30	30	110	12.5	34.5
AIRS 530330080881011	89	83	80	87	339	9.1	25.0
Composite[1]	91	91	92	92	366	10.3	27.8

*All concentrations are in $\mu g/m^3$; includes King County.

1. The number of days at the composite monitor is the number of days on which at least one of the monitors reported.

Exhibit A.10. Monitor Locations Proposed for Use in $PM_{2.5}$ Risk Analyses in Los Angeles.



Monitor Legend

1	AIRS-060372811011
2	AIRS-061110011011
3	AIRS-060371201011
4	AIRS-060371301011
5	AIRS-060371601011
6	AIRS-060372005011
7	AIRS-060374002811011

2000 Population: 9,519,338

N

Exhibit A.11. Monitor Locations Proposed for Use for $PM_{2.5}$ Analyses in Philadelphia.

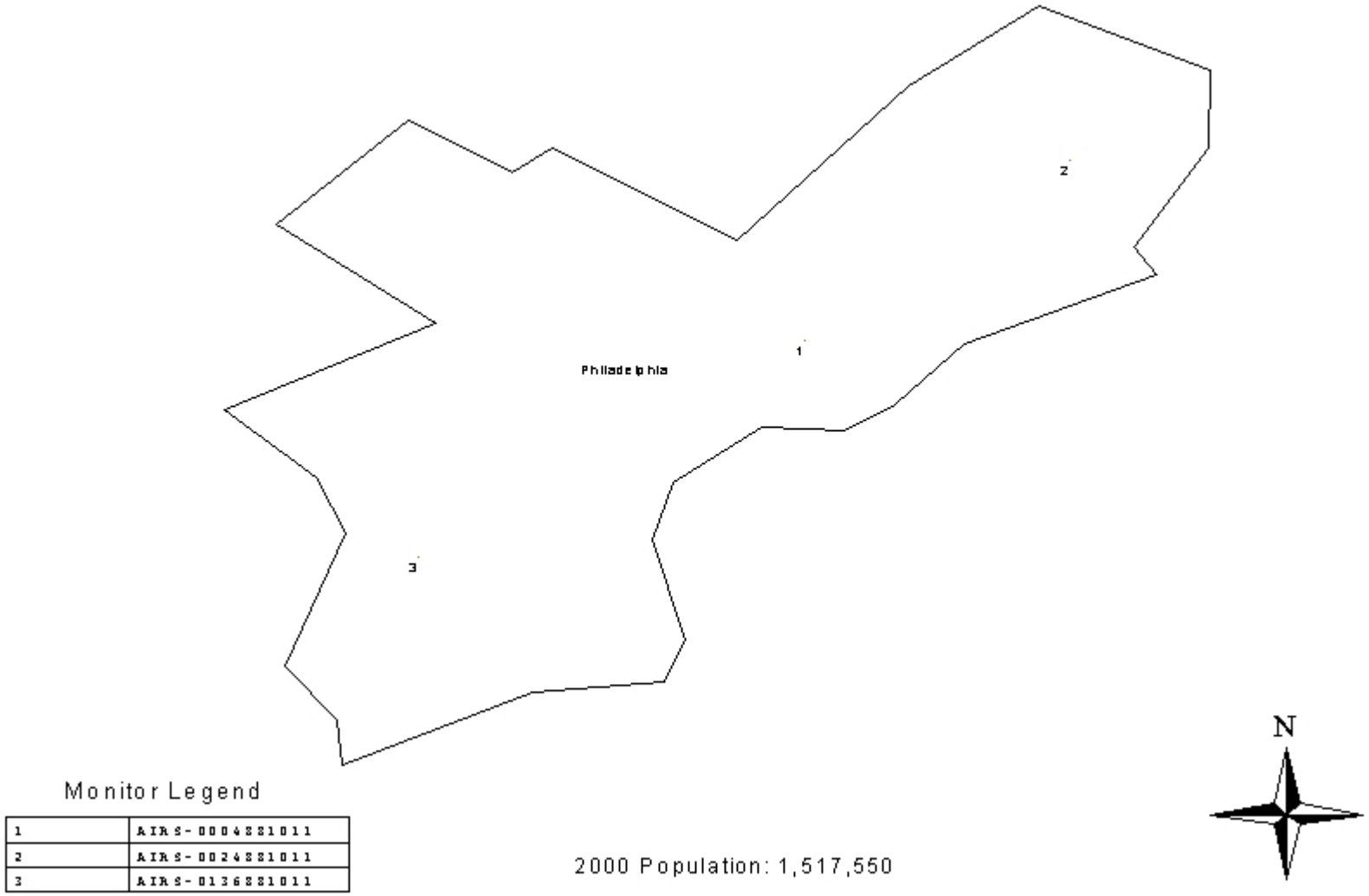


Exhibit A.12. Monitor Locations Proposed for Use in $PM_{2.5}$ Risk Analyses in Philadelphia Region.

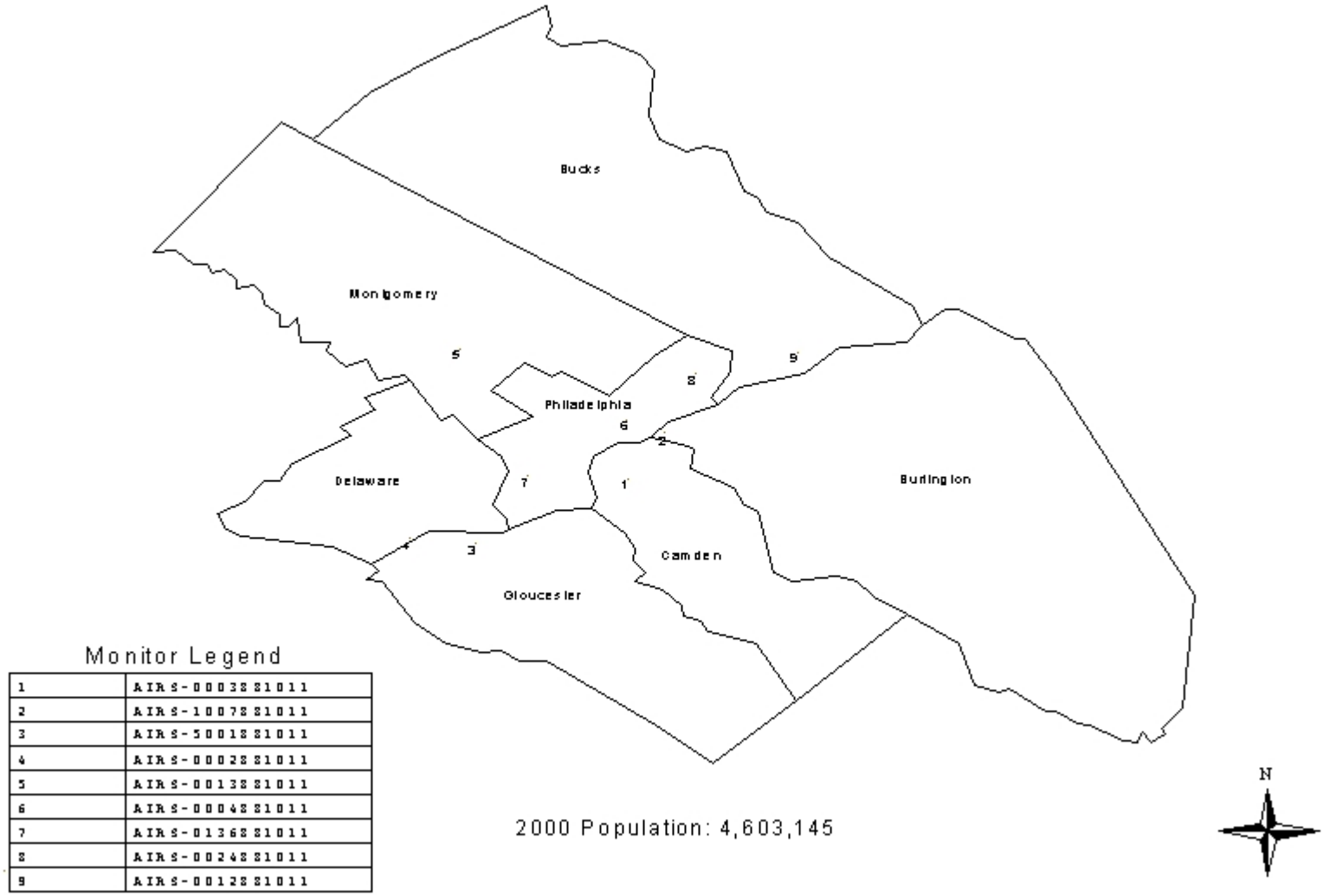
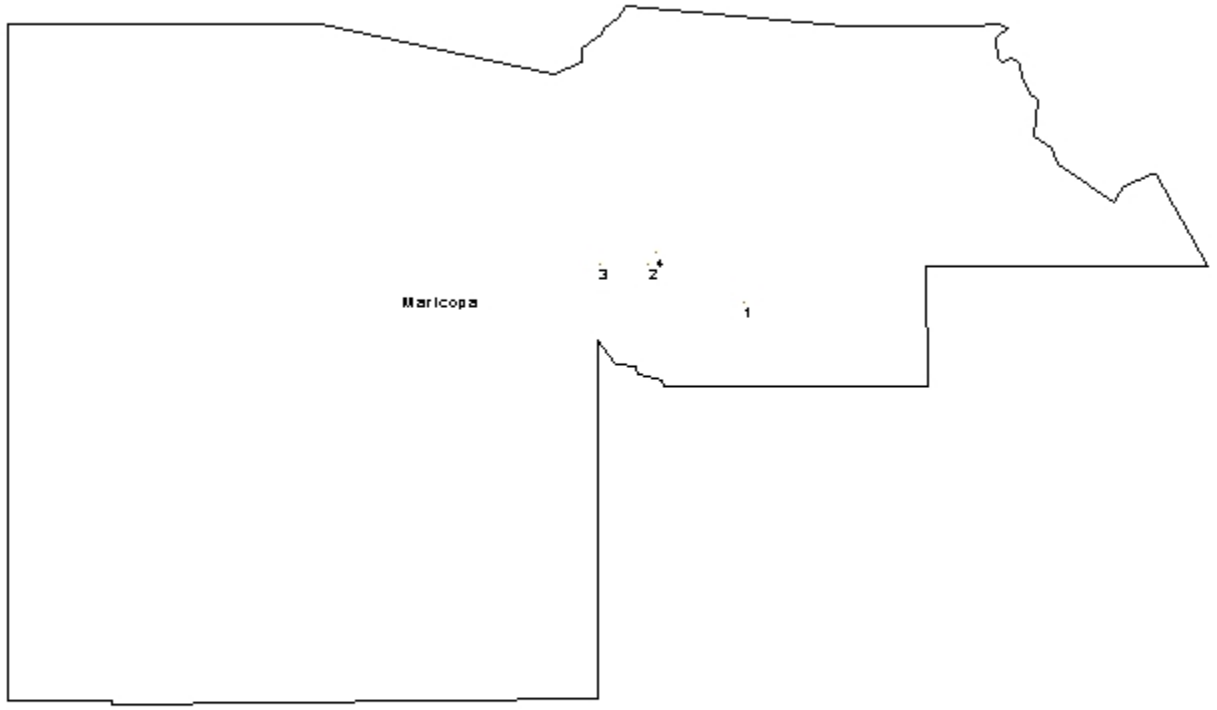


Exhibit A.13. Monitor Locations Proposed for Use in $PM_{2.5}$ Risk Analyses in Phoenix.



Monitor Legend

1	AIRS-999083101 1
2	AIRS-999183101 1
3	AIRS-999283101 1
4	AIRS-999783101 1

2000 Population: 3,072,149

N

Exhibit A.14. Monitor Locations Proposed for Use in $PM_{2.5}$ Risk Analyses in San Jose.

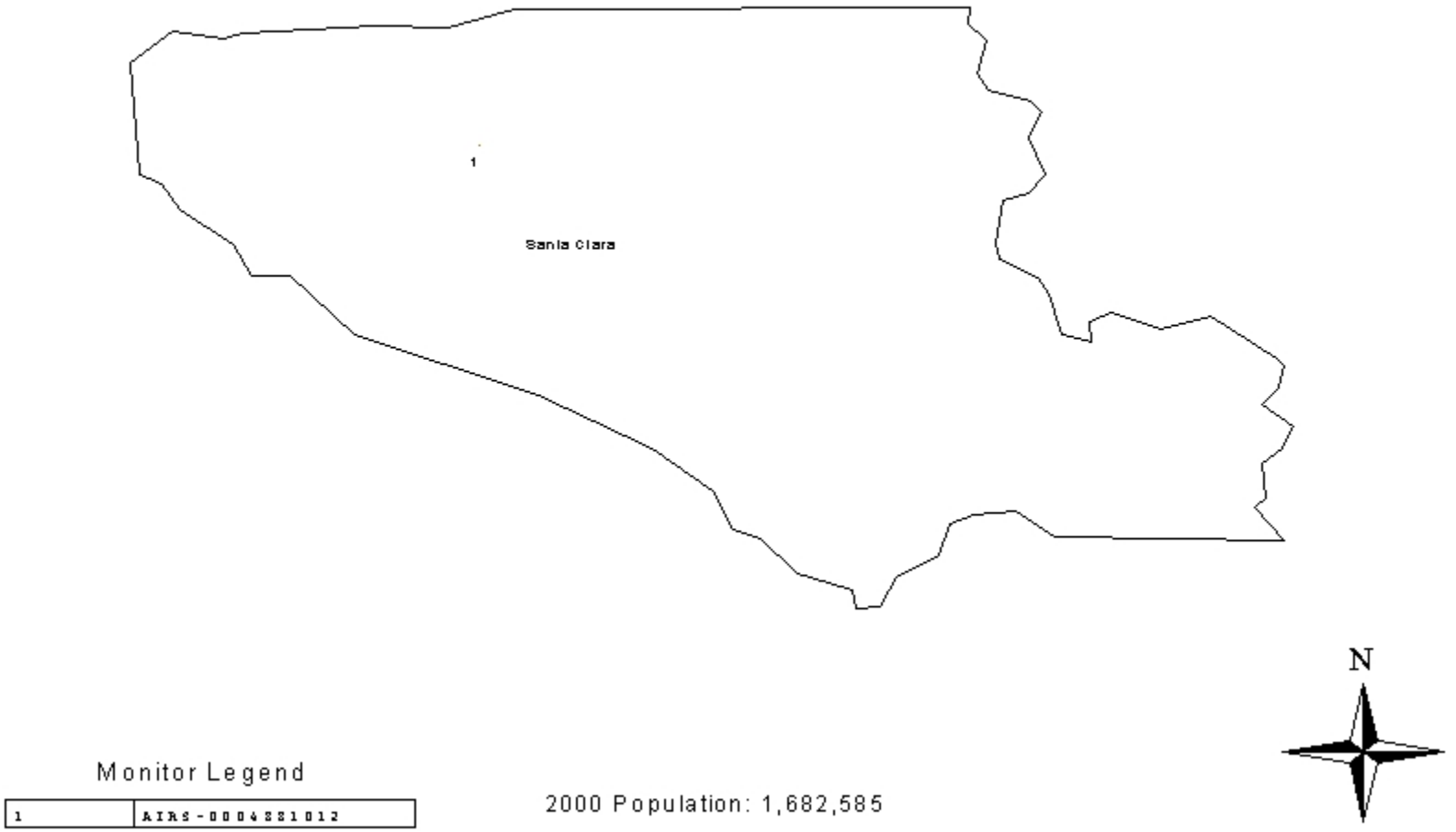
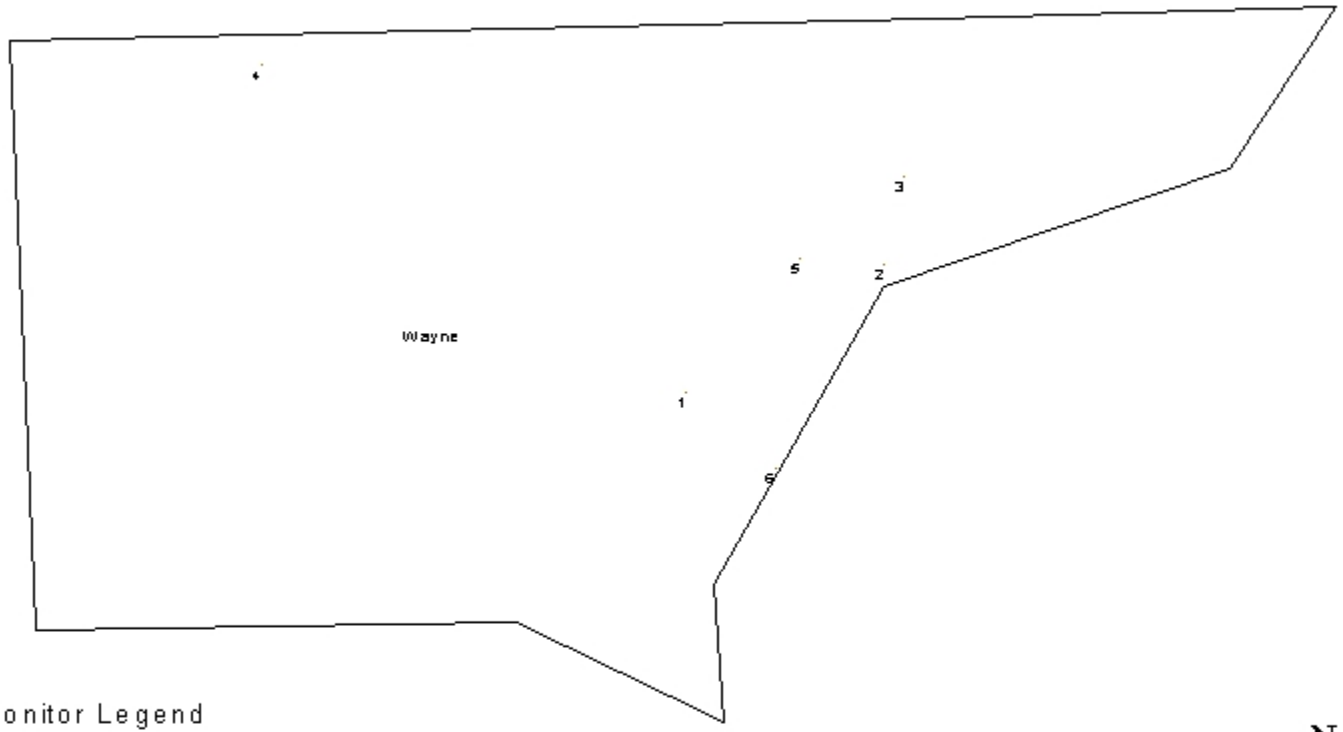


Exhibit A.15. Monitor Locations Proposed for Use in $PM_{2.5}$ Risk Analyses in Boston.

Exhibit A.16. Monitor Locations Proposed for Use in $PM_{2.5}$ Risk Analyses in Detroit.



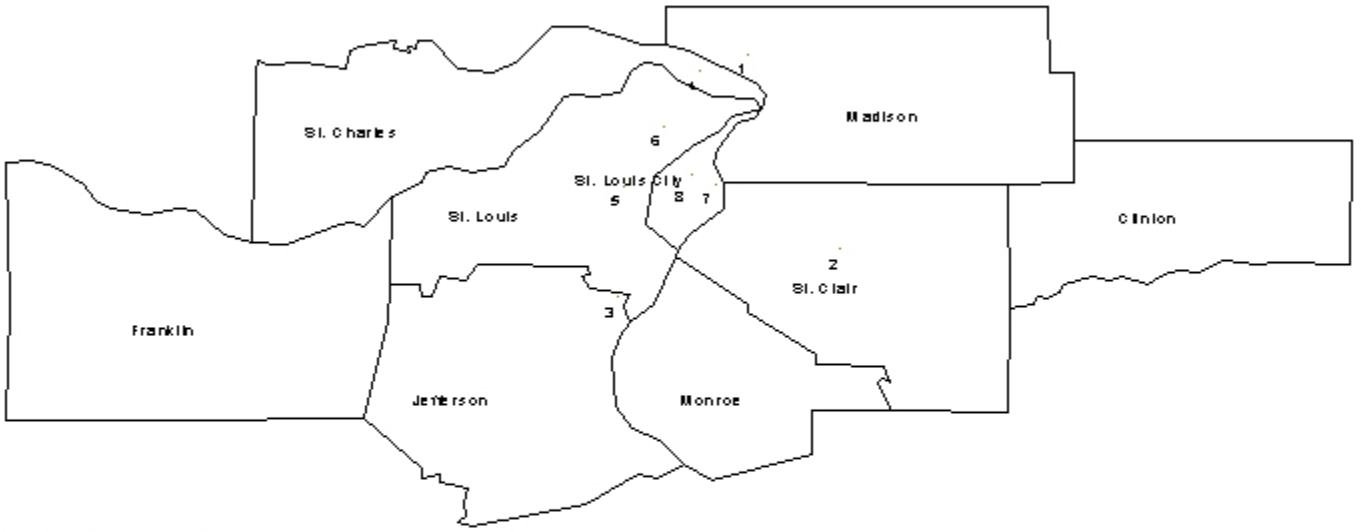
Monitor Legend

1	AIRS-000183 1011
2	AIRS-001583 1011
3	AIRS-001683 1011
4	AIRS-002583 1011
5	AIRS-003383 1011
6	AIRS-003683 1011

2000 Population: 2,061,162

N

Exhibit A.17. Monitor Locations Proposed for Use in $PM_{2.5}$ Risk Analyses in St. Louis.



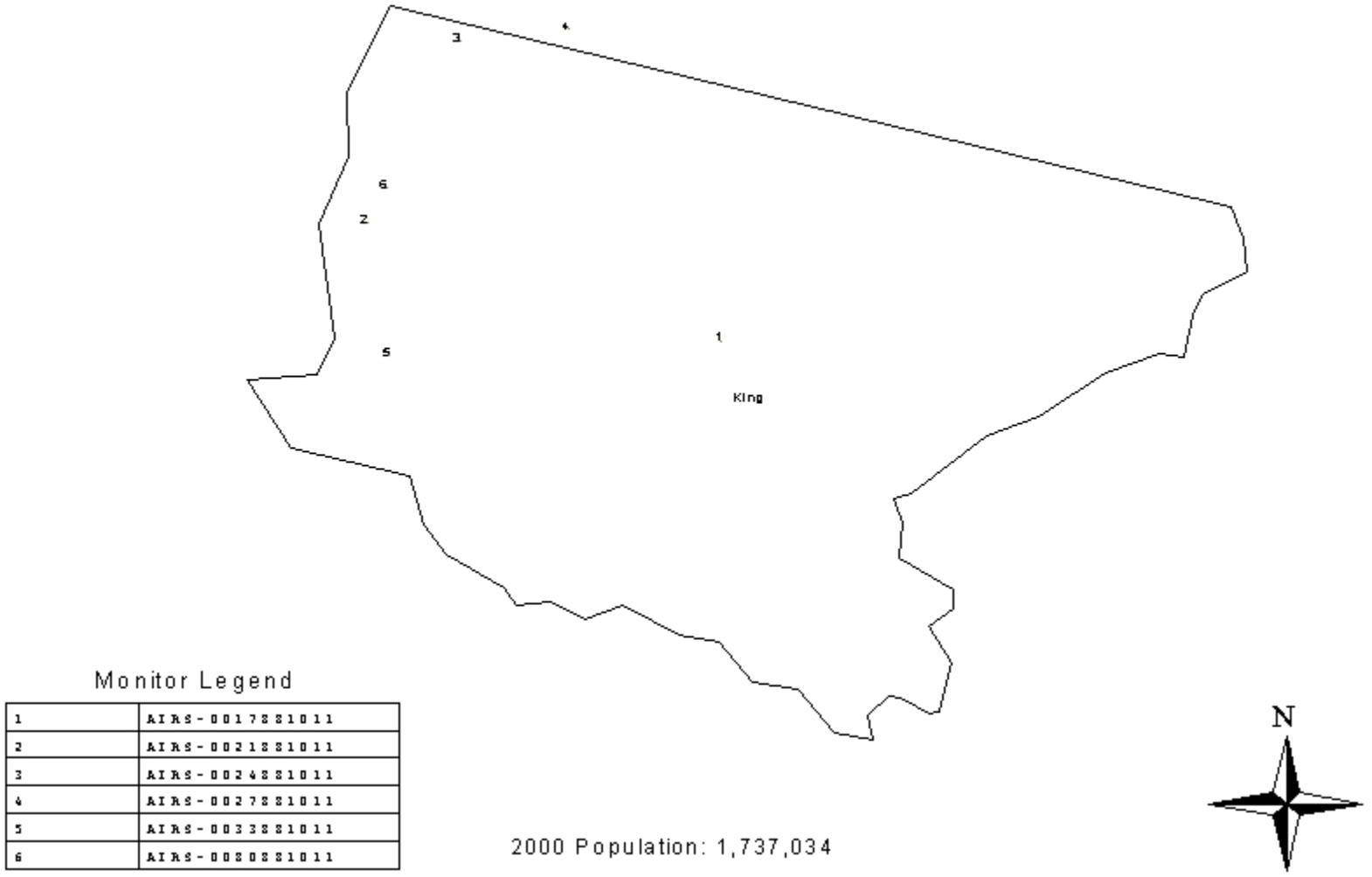
Monitor Legend

1	AIRS-2909881011
2	AIRS-40018811011
3	AIRS-00128811011
4	AIRS-10028811011
5	AIRS-29038811011
6	AIRS-50018811011
7	AIRS-00858811011
8	AIRS-00868811011

2000 Population: 2,518,470

N

Exhibit A.18. Monitor Locations Proposed for Use in $PM_{2.5}$ Risk Analyses in Seattle.



Appendix B. Linear Trends in Historical $PM_{2.5}$ Data in Philadelphia and Los Angeles

Abt

memorandum

Environmental Research Area
4800 Montgomery Lane, Suite 600 ■ Bethesda, MD 20814-5341 ■ (301) 913-0500

Abt Associates Inc.

Date November 26, 2001

To Harvey Richmond, U.S. EPA/OAQPS

From Ellen Post, Abt Associates Inc.

Subject **Linear Trends in Historical $PM_{2.5}$ Data in Philadelphia and Los Angeles**

The method used to simulate just meeting a standard in the 1995/96 PM risk analysis and proposed for the current risk analysis is to "roll back" the anthropogenic portion of PM levels (i.e., the portion above background level) by the same percentage on each day. This method assumes that, all else held constant:

$$(y_i - B) = \beta^*(x_i - B) \quad ,$$

where[1]

- x_i is the ith $PM_{2.5}$ concentration in a location before the standard is met,
- y_i is the ith $PM_{2.5}$ concentration in that location when the standard is just met,
- B is the background concentration in that location, and
- $\beta < 1$.

We don't have data on $PM_{2.5}$ concentrations in any location before and after the $PM_{2.5}$ standards have just been met, so we cannot directly test whether this "rollback" assumption accurately models how $PM_{2.5}$ concentrations would change if a standard were just met. We can, however, examine historical changes in $PM_{2.5}$ concentrations for any location for which we have sufficient data to determine if the proportional rollback model is consistent with these historical changes. We currently have sufficient data in each of two locations, Philadelphia and Los Angeles, to compare the distribution of daily $PM_{2.5}$ concentrations in the year 2000 with the distribution in an earlier year. In each location, we compared

---

[1] We first examined the plausibility of this assumption in preparation for the PM risk analyses carried out in 1995/1996. At that time, we examined pairs of years of $PM_{2.5}$ data in several locations, but none of the data reflected efforts to meet $PM_{2.5}$ standards, because this exercise (and the data it used) preceded the setting of $PM_{2.5}$ standards. That investigation, however, found that the change in the distribution of $PM_{2.5}$ concentrations from one year to another year in the same location tended to be linear. This is described in Section 8.2 of Abt Associates Inc., 1996. "A Particulate Matter Risk Assessment for Philadelphia and Los Angeles."

the two distributions to see if the change was well described as proportional. The method and results are described below.

In Philadelphia we have 353 days of observations in a year which crosses calendar years 1992 and 1993, and 296 days of observations in the year 2000. In Los Angeles we have 214 days of observations in 1995 and 357 days of observations in 2000. We first grouped the $PM_{2.5}$ concentrations in each distribution into deciles and averaged the concentrations within each decile.[2] These average concentrations within deciles are shown in Exhibit B.1 and in graph form in Exhibits B.2 and B.3, for Philadelphia and Los Angeles, respectively.

Exhibit B.1. Average $PM_{2.5}$ Concentrations ($\mu g/m^3$) in Each Decile of Earlier Year and Year 2000 Distributions at Composite Monitors in Philadelphia and Los Angeles*

Decile*	Philadelphia		Los Angeles	
	1992/93	2000	1995	2000
1	5.91	4.62	10.02	6.67
2	7.94	6.58	14.62	10.19
3	9.71	8.82	18.50	12.39
4	11.19	10.25	21.06	14.59
5	13.07	12.07	24.19	16.59
6	14.87	13.72	28.40	18.55
7	17.23	16.01	32.96	21.27
8	20.67	19.4	39.72	24.22
9	25.34	23.77	54.77	28.27
10	37.90	32.58	87.12	50.50

*The first decile is the tenth percentile, the second decile is the twentieth percentile, and so on. The average concentration in the nth decile is the average of those values that are greater than the (n-1)st decile point and less than or equal to the nth decile point.

[2] We considered using the decile points themselves rather than the averages within deciles. However, the decile points would be expected to be less stable from one year to another than the averages of the concentrations within deciles. A comparison of the averages within deciles from one year to another is therefore likely to give a more accurate picture of how the distribution has changed from one year to another. This is the method that was used in the earlier comparison for the 1995/96 PM risk analysis.

Exhibit B.2

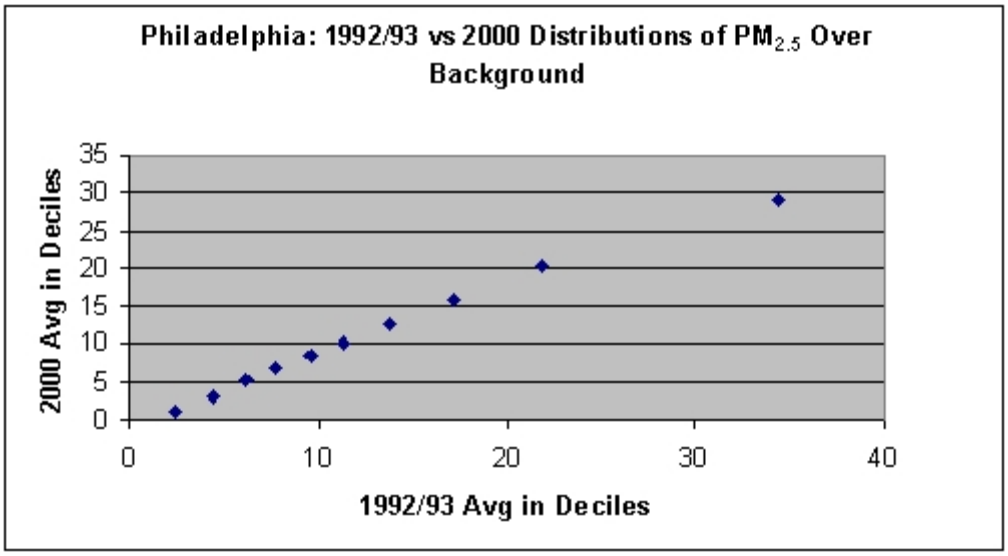
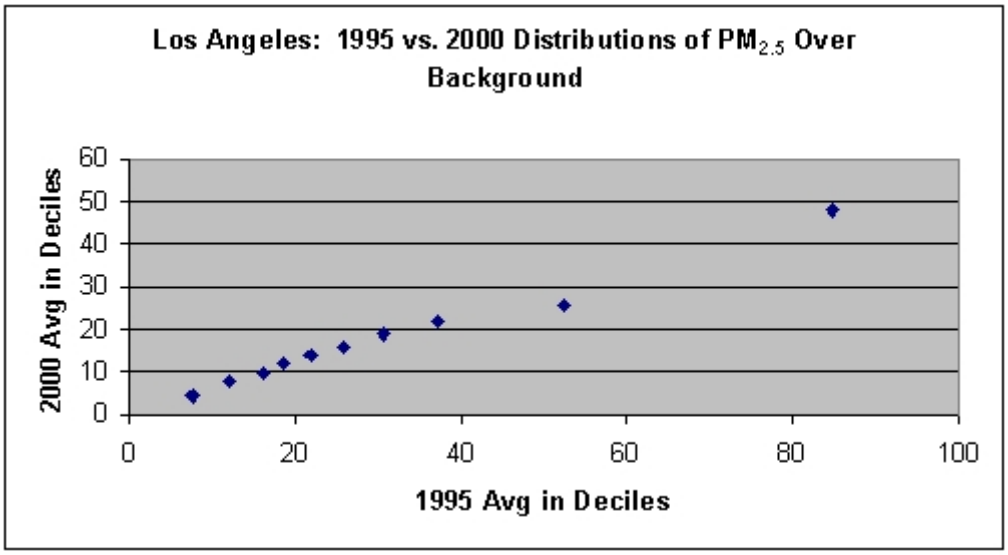


Exhibit B.3



To test the proportional rollback hypothesis we estimated the following regression equation separately for Philadelphia and for Los Angeles:

$$(y_i - B) = \alpha + \beta^*(x_i - B) + \varepsilon_i$$

where now,

- y_i is the average $PM_{2.5}$ concentration in the ith decile of the distribution of $PM_{2.5}$ concentrations in the location in the year 2000,
- x_i is the average $PM_{2.5}$ concentration in the ith decile of the distribution of $PM_{2.5}$ concentrations in that location in an earlier year (1995 for Los Angeles and 1992/93 for Philadelphia),
- B is the background concentration in that location (2.5 $\mu g/m^3$ in Los Angeles and 3.5 $\mu g/m^3$ in Philadelphia), and
- ε_i is an error term.

If the change in $PM_{2.5}$ concentrations from the earlier year to the year 2000 is consistent with a proportional rollback model, we would expect

- the linear fit to be good,
- the slope (β) to be statistically significant and less than one, and
- the intercept (α) to be not statistically significantly different from zero

The results of the regressions in Philadelphia and Los Angeles do support the hypothesis underlying the proportional rollback method, as shown in Exhibit B.4. In both cases, the linear fit is very good ($R^2 = 0.992$ in Philadelphia and 0.986 in Los Angeles), the slopes are highly statistically significant and less than 1.0, and the intercepts are not significantly different from zero. This supports the hypothesis that, at least in these two locations, the change in daily $PM_{2.5}$ concentrations that would result if a $PM_{2.5}$ standard were just met is reasonably modeled as a proportional rollback.

Exhibit B.4. Results of Regressions of Year 2000 Average $PM_{2.5}$ Concentrations over Background on Earlier Year Average $PM_{2.5}$ Concentrations over Background.

	Philadelphia	Los Angeles
Intercept	-0.136 (p=0.76)	1.387 (p=0.146)
Slope	0.886 ($p=9.56 \times 10^{-10}$)	0.537 ($p=1.16 \times 10^{-8}$)
R^2	0.992	0.986

Appendix C. Air Quality, Health Studies, and Concentration-Response Relationships

This Appendix summarizes the $PM_{2.5}$ air quality information and health effects studies that were used as a basis for developing the proposed list of health effect endpoints and urban locations to include in the $PM_{2.5}$ risk analyses. Given the large number of endpoints and studies addressing $PM_{2.5}$ effects, EPA is proposing to include in the quantitative risk analyses only the better understood (in terms of health consequences) health endpoint categories for which the weight of the evidence supports the existence of a relationship between $PM_{2.5}$ and the effect category and only those categories which included studies that directly measured fine fraction PM using $PM_{2.5}$ or $PM_{2.1}$ as the indicator.

Exhibit C.1 presents a summary of the U.S. and Canadian studies identified in the draft PM CD (U.S. EPA, 2001a) that report effect estimates for short-term exposure mortality associated with $PM_{2.5}$ (or $PM_{2.1}$). The U.S. locations in this exhibit represent the candidate pool of possible locations to include in the $PM_{2.5}$ risk analyses for short-term exposure mortality. Generally studies identified in Table 9-3 of the draft PM CD were included, with the exception of studies that did not directly measure either $PM_{2.5}$ or $PM_{2.1}$. Effect estimates for $PM_{2.5}$ and summary information about $PM_{2.5}$ ambient concentrations measured in these studies is provided in Exhibit C.1. The last column in Exhibit C.1 is the natural log of the product of mortality rate and number of days in the study, which, as discussed below and in the 2001 draft PM CD and preliminary PM SP, is a surrogate measure of the relative statistical power of the study to detect health effects associated with air pollutants.

As discussed in Section 3.2.1 of this report, several considerations were used in deriving the list of proposed urban areas to be included in the PM risk analyses examining short-term exposure mortality. In narrowing the list from the larger candidate pool, we first considered the statistical power of the studies that estimated $PM_{2.5}$ short-term exposure mortality C-R functions in those locations. In general, the power of a study increases as the number of its observations increases. The number of observations depends not only on the number of days on which mortality counts were obtained, but also on the size of the mortality counts. The 2001 draft PM CD uses the natural logarithm of the mortality-days (i.e., the natural log of the product of the number of study days and the average number of deaths per day) as a surrogate or indicator reflecting the power of short-term exposure mortality epidemiological studies. As stated in the 2001 draft PM CD (pp.6-260, 6-263), "the more the mortality-day observations, the narrower the 95% confidence intervals and the more precise the effects estimates (with nearly all these for cities with $\geq$ log 9 mortality-days being positive and many statistically significant at $p \leq 0.05$)." OAQPS proposes to consider for inclusion in the risk analyses only those locations in which studies with greater precision were conducted as indicated by having a natural log of total non-accidental mortality-days greater than or equal to 9.0.

We next considered which of those study locations have sufficient $PM_{2.5}$ monitoring data to support a risk analysis. Air quality data were obtained from EPA's Aerometric Information

Retrieval System (AIRS) with all observations with validation flags[1] 1, 2, 3, or 4 deleted for the year 2000 (or 1999 where there was not adequate data for an area in the year 2000) for cities in which candidate $PM_{2.5}$ epidemiological studies had been conducted based on the studies summarized in Exhibits C.1 and C.2. Exhibit C.3 shows the monitor-specific minimum number of observations per quarter and the number of observations per year for all of the U.S. locations for the studies summarized in Exhibits C.1 and C.2 that met the cutoff for statistical power described above. All federal reference method monitors in each of the locations are presented as well as the minimum number of monitoring days per quarter and the total count per year. We propose to exclude those monitors which are identified in AIRS as targeting "highest concentration" as their monitoring objective, which are generally located in either an "industrial" or "commercial" land use area based on the information from AIRS. For any monitor which is not thus excluded but is not specifically identified as "population-oriented," Abt plans to evaluate the extent to which the daily $PM_{2.5}$ concentrations at the monitor are correlated with those at the other monitors in the urban area to which it belongs. In cases where there is a very low correlation, we propose to drop the monitor from the analysis and to exclude the population living near it.

An urban area was considered to have sufficiently complete air quality data if it had at least one monitor at which there were at least 11 observations in each quarter and at least 122 observations per year (equivalent to at least 1 in 3 day monitoring). Using that completeness criterion, seven areas listed in Exhibit C.3 in which short-term mortality C-R functions were estimated had sufficient year 2000 data (including Philadelphia region). One additional area (Seattle), in which a hospitalization C-R function was estimated, also had sufficient year 2000 data.. Although Phoenix did not have sufficient data in 2000, it did have sufficient data in 1999. The studies in bold typeface in Exhibit C.1 indicate the non-accidental short-term exposure mortality study locations which had sufficient $PM_{2.5}$ air quality data and which were judged as having relatively higher statistical power.

Once the criteria for inclusion of a study location are met, all monitors with at least 11 observations per quarter at that location will be used. The cutoff of 11 observations per quarter is based on EPA guidance on measuring attainment of the daily and annual particulate matter standards outlined in Appendix N of the July 18, 1997 Federal Register Notice (available on the web at www.epa.gov/ttn/oarpg/t1pfpr.html). The guidance requires that at least 75 percent of the scheduled sampling days for each quarter have valid data. Based on a one in six day sampling protocol, the minimum required number of observations would be 11 per quarter. Those monitors that meet this criterion and will be used in the risk analyses are indicated in bold typeface in Exhibit C.3.

Most of the urban locations in which C-R functions were estimated for health endpoints other than mortality are included in the set of locations available for mortality (see Exhibit C.2

[1]Validation flags are placed in AIRS by the State and/or local air pollution agencies and relate to the potential validity of the data reported. Data with these flags were excluded from the analyses because final determinations had not yet been made regarding the validity of this data.

for hospital admission, emergency room visit, and respiratory symptom and illness studies). A primary consideration in selecting urban locations for these other health endpoints, as with the risk analyses for mortality, is that the assessment locations be the same as or close to the study locations for which estimated C-R functions were reported. In addition, studies with relatively higher statistical power to detect relatively small but real population effects are preferable. As with mortality, another consideration is the availability of recent and adequate $PM_{2.5}$ air quality data.[2] Finally, for the hospital admission effect category, the availability of baseline incidence data is an additional consideration in selecting urban locations for the risk analyses. Data on hospital admissions for recent years, specific to International Classification of Disease (ICD) codes, are available in some cities but not others. Based on all of the above considerations, the proposed locations for conducting the $PM_{2.5}$ risk analyses for hospital admissions and emergency visits were selected and they are indicated in bold typeface in Exhibit C.2. In addition to those studies indicated in bold, we propose to use the estimated C-R relationships reported in Schwartz and Neas (2000), a study conducted across several cities, for the respiratory symptom endpoints.

Exhibit C.4 presents a summary of the U.S. and Canadian studies identified in the draft PM CD (U.S. EPA, 2001a) that report effect estimates for long-term exposure mortality associated with $PM_{2.5}$. The studies in bold, proposed for use in this risk analysis, are re-analyses of the Dockery et al. (1993) and Pope et al. (1995) long-term exposure studies of the association between annual measures of $PM_{2.5}$ and all-cause mortality. Since these studies included multiple locations in the U.S., we will apply the C-R functions to each of the eight areas that have been proposed for the short-term exposure mortality analyses.

[2]As noted earlier, for any monitor which is not excluded because it has a "highest concentration" monitoring objective but is not specifically identified as "population-oriented," Abt plans to evaluate the extent to which the daily $PM_{2.5}$ concentrations at the monitor are correlated with those at the other monitors in the urban area to which it belongs. In cases where there is a very low correlation, we propose to drop the monitor from the analysis and to exclude the population living near it.

Exhibit C.1. Estimated Increased Mortality per Increments in 24-h Concentrations of $PM_{2.5}$ from U.S. and Canadian Studies.

Reference, Study Location *	% increase (95% CI) per 25 μg/m^3 $PM_{2.5}$ Increase	$PM_{2.5}$ Mean (Range) Levels Reported**	ln mortality-days
Total (nonaccidental) Mortality			
Schwartz et al., 1996 Boston, MA	*5.59 (3.80, 7.42)*	*$PM_{2.5}$ 15.7 (SD 9.2)*	*11.1*
Schwartz et al., 1996 Knoxville, TN	*3.54 (0.52, 6.65)*	*$PM_{2.5}$ 20.8 (SD 9.6)*	*10.0*
Schwartz et al., 1996 St. Louis, MO	*2.77 (1.13, 4.44)*	*$PM_{2.5}$ 18.7 (SD 10.5)*	*11.1*
Schwartz et al., 1996 Steubenville, OH	*2.52 (-0.24, 5.35)*	*$PM_{2.5}$ 29.6 (SD 21.9)*	*8.6*
Schwartz et al., 1996 Portage, WI	*3.03 (-0.84, 7.05)*	*$PM_{2.5}$ 11.2 (SD 7.8)*	*9.9*
Schwartz et al., 1996 Topeka, KS	*2.01 (-4.83, 9.35)*	*$PM_{2.5}$ 12.2 (SD 7.4)*	*8.8*
Schwartz et al., 1996 6 Cities, Overall	*3.79 (2.77, 4.82)*	*$PM_{2.5}$ means 11.2-29.6*	*12.2*
Burnett et al., 1998 Toronto, Canada	4.79 (3.26, 6.34)	$PM_{2.5}$ 18.0 (8, 90)	12.3
Burnett et al., 2000 8 Canadian Cities	3.03 (1.10, 4.99)	$PM_{2.5}$ 13.3 (max 86)	11.7
Fairley, 1999 San Jose, CA	8 ($p<0.01$)	$PM_{2.5}$ 13 (2, 105)	9.0
Goldberg et al., 2000 Montreal, Canada	5.81 (3.36, 8.32)	$PM_{2.5}$ 17.6 (4.6, 71.7)	11.9
Lipfert et al., 2000 Philadelphia, PA	4.21 ($p<0.055$)	$PM_{2.5}$ 17.28 (-0.6, 72.6)	11.8
Lippmann et al., 2000 Detroit, MI	3.10 (-0.63, 6.98)	$PM_{2.5}$ 18 (6, 86) mean (5%, 95%)	9.8
Mar et al., 2000 Phoenix, AZ	5.98 (-1.34, 13.85)	$PM_{2.5}$ 13.0 (0, 42)	9.1
Moolgavkar, 2000a Los Angeles, CA	0.6 ($p>0.05$, from figure)	$PM_{2.5}$ 22 (4, 86)	13.1
Schwartz, 2000c Boston, MA	5.33 (1.81, 8.98)	$PM_{2.5}$ 15.6 (±9.2)	12.1
Tsai et al., 2000 Newark, NJ	4.34 (2.82, 5.89)	$PM_{2.5}$ 42.1 (SD 22.0)	8.7
Tsai et al., 2000 Camden, NJ	5.65 (0.11, 11.51)	$PM_{2.5}$ 39.9 (SD 18.0)	7.4

Reference, Study Location *	% increase (95% CI) per 25 μg/m^3 $PM_{2.5}$ Increase	$PM_{2.5}$ Mean (Range) Levels Reported**	ln mortality-days
Tsai et al., 2000 Elizabeth, NJ	1.77 (-5.44, 9.53)	$PM_{2.5}$ 37.1 (SD 19.8)	7.6
Cause-Specific Mortality			
Cardiorespiratory:			
Tsai et al., 2000 Newark, NJ	5.13 (3.09, 7.21)	$PM_{2.5}$ 42.1 (SD 22.0)	8.1
Tsai et al., 2000 Camden, NJ	6.18 (0.61, 12.06)	$PM_{2.5}$ 39.9 (SD 18.0)	6.8
Tsai et al., 2000 Elizabeth, NJ	2.28 (-4.97, 10.07)	$PM_{2.5}$ 37.1 (SD 19.8)	7.0
Total Cardiovascular:			
Fairley, 1999 Santa Clara County (San Jose), CA	6.2 (p>0.05)	$PM_{2.5}$ 13 (2, 105)	8.2
Goldberg et al., 2000 Montreal, Canada	3.48 (-0.16, 7.26)	$PM_{2.5}$ 17.6 (4.6, 71.7)	11.0
Lipfert et al., 2000 Philadelphia, PA (7-county area)	10.26 (p<0.055)	$PM_{2.5}$ 17.28 (-0.6, 72.6)	11.0
Lippmann et al., 2000 Detroit, MI	3.17 (-2.29, 8.94)	$PM_{2.5}$ 18 (6, 86) mean (10%, 90%)	9.1
Mar et al., 2000 Phoenix, AZ	18.68 (5.72, 33.23)	$PM_{2.5}$ 13.0 (0, 42)	8.3
Moolgavkar, 2000a Los Angeles, CA	2.59 (0.38, 4.85)	$PM_{2.5}$ median 22 (4, 86)	12.1
Coronary Artery Disease:			
Goldberg et al., 2000 Montreal, Canada	4.48 (-0.31, 9.51)	$PM_{2.5}$ 17.6 (4.6, 71.7)	11.0
Total Respiratory:			
Fairley, 1999 Santa Clara County (San Jose), CA	11.5 (p>0.05)	$PM_{2.5}$ 13 (2, 105)	6.9
Goldberg et al., 2000 Montreal, Canada	21.6 (13.0, 31.0)	$PM_{2.5}$ 17.6 (4.6, 71.7)	9.3
Lipfert et al., 2000 Philadelphia, PA (7-county area)	0.66 (p>0.055)	$PM_{2.5}$ 17.28 (-0.6, 72.6)	9.4
Lippmann et al., 2000 Detroit, MI	2.28 (-10.31, 16.63)	$PM_{2.5}$ 18 (6, 86) mean (10%, 90%)	7.2

Reference, Study Location *	% increase (95% CI) per 25 $\mu g/m^3$ $PM_{2.5}$ Increase	$PM_{2.5}$ Mean (Range) Levels Reported**	ln mortality-days
Respiratory (COPD & asthma):			
Moolgavkar, 2000a Los Angeles, CA	2.67 (-3.38, 9.10)	$PM_{2.5}$ 22 (4, 86)	9.9

* Studies in italics available in 1996 CD. Studies in bold indicate studies that support choice of proposed locations to include in $PM_{2.5}$ risk analyses based on measure of statistical power of study and availability of sufficient recent $PM_{2.5}$ air quality data.

** Mean (minimum, maximum) 24-h PM level shown in parentheses unless otherwise noted.

Exhibit C.2. Estimated Respiratory and Cardiovascular Morbidity Effects per Increments in 24-h Concentrations of $PM_{2.5}$ from U.S. and Canadian Studies.

Reference, Study Location*	% increase (95% CI) per 25 µg/m^3 $PM_{2.5}$ Increase	$PM_{2.5}$ Mean (Range) Levels Reported**	ln admission-days (or Emergency Room Visits)
Increased Admission to Hospital (or Emergency Room Visit)****			
Total Respiratory:			
Thurston et al., 1994 *Toronto, Canada*	*15.00 (1.97, 28.03)*	*Summers 1986-1988* *$PM_{2.5}$ 15.8-22.3 (max 66.0)*	*7.4*
Burnett et al., 1997 Toronto, CAN (all ages)	8.61 (3.39, 14.08)	Summers 1992-1994 $PM_{2.5}$ 16.8 (1, 66)	9.1
Delfino et al., 1997 Montreal, CAN (>64 years)	23.88 (4.94, 42.83)	summer 93 $PM_{2.5}$ 12.2 (max 31)	7.8
Delfino et al., 1998 Montreal, CAN (>64 years)	13.17 (-0.22, 26.57)	$PM_{2.5}$ 18.6 (SD 9.3)	7.5
Stieb et al., 2000**** St. John, CAN (all ages)	5.69 (0.61, 11.03)	$PM_{2.5}$ 8.5 (max 53.2)	9.5
Pneumonia:			
Lippmann et al., 2000 **Detroit, MI (>65 years)**	12.5 (3.7, 22.1)	$PM_{2.5}$ 18 (6, 86)	8.7
Respiratory infections:			
Burnett et al., 1999 Toronto, CAN (all ages)	10.77 (7.18, 14.47)	$PM_{2.5}$ 18.0 (max 90)	11.2
COPD:			
Tolbert et al., 2000a**** Atlanta, GA (all ages)	12.44 (-7.89, 37.24)	$PM_{2.5}$ 19.4 (SD 9.35)	8.1
Lippmann et al., 2000 **Detroit, MI (>65 years)**	5.49 (-4.72, 16.80)	$PM_{2.5}$ 18 (6, 86)	8.3
Moolgavkar et al., 2000 **King County WA (all ages)**	6.4 (0.9, 12.1)	$PM_{2.5}$ 18.1 (3, 96)	8.9
Moolgavkar, 2000c **Los Angeles, CA (>65 years)**	5.1 (0.9, 9.41)	$PM_{2.5}$ median 224, 86)	11.1
Burnett et al., 1999 Toronto, CAN (all ages)	4.78 (-0.17, 9.98)	$PM_{2.5}$ 18.0 (max 90)	10.2
Asthma:			

Tolbert et al., 2000a**** Atlanta, GA (all ages)	2.27 (-14.79, 22.74)	$PM_{2.5}$ 19.4 (SD 9.35)	8.6
Burnett et al., 1999 Toronto, CAN (all ages)	6.44 (2.47, 10.57)	$PM_{2.5}$ 18.0 (max 90)	11.0
Total Cardiovascular:			
Moolgavkar, 2000b **Los Angeles, CA (>65 years)**	(65+) 4.30 (2.52, 6.11) (<65) 3.54 (1.83, 5.27)	$PM_{2.5}$ median 22 (4, 86)	13.2
Tolbert et al., 2000a**** Atlanta, GA (all ages)	6.11 (-3.08, 16.17)	$PM_{2.5}$ 19.4 (SD 9.35)	9.7
Stieb et al., 2000**** St. John, CAN (all ages)	15.11 (0.61, 11.03)	$PM_{2.5}$ 8.5 (max 53.2)	8.4
Burnett et al., 1997 Toronto, CAN (all ages)	7.18 (-0.61, 15.60)	summers 1992-1994 $PM_{2.5}$ 16.8 (1, 66)	9.7

Reference, Study Location*	% increase (95% CI) per 25 $\mu g/m^3$ $PM_{2.5}$ Increase	$PM_{2.5}$ Mean (Range) Levels Reported**	ln admission-days (or Emergency Room Visits)
Ischemic Heart Disease:			
Lippmann et al., 2000 **Detroit, MI (>65 years)**	4.33 (-1.39, 10.39)	$PM_{2.5}$ 18 (6, 86)	9.3
Burnett et al., 1999 Toronto, CAN (all ages)	8.05 (5.38, 10.78)	$PM_{2.5}$ 18.0 (max 90)	11.8
Heart Failure:			
Lippmann et al., 2000 **Detroit, MI (>65 years)**	9.06 (2.36, 16.19)	$PM_{2.5}$ 18 (6, 86)	9.0
Burnett et al., 1999 Toronto, CAN (all ages)	6.59 (2.50, 10.83)	$PM_{2.5}$ 18.0 (max 90)	10.8
Increased Respiratory Symptoms	Odds Ratio (95% CI) for 25 ug/m^3 increase in $PM_{2.5}$	$PM_{2.5}$ Mean (Range) Levels Reported**	
Schwartz et al., 1994 *6 U.S. cities**** *(children, cough)*	*1.24 (1.00, 1.54)*	*$PM_{2.5}$ median 18.0 (max 86)*	
Schwartz et al., 1994 *6 U.S. cities**** *(children, lower respiratory symptoms)*	*1.58 (1.18, 2.10)*	*$PM_{2.5}$ median 18.0 (max 86)*	

Increased Respiratory Symptoms	Odds Ratio (95% CI) for 25 ug/m^3 increase in $PM_{2.5}$	$PM_{2.5}$ Mean (Range) Levels Reported**
Neas et al., 1995 Uniontown, PA (children, cough)	*2.45 (1.29, 4.64)*	*$PM_{2.5}$ 24.5 (max 88.1)*
Neas et al., 1996 State College, PA (children, cough)	1.48 (1.17, 1.88) (1-d)	$PM_{2.1}$ 23.5 (max 85.8)
Neas et al., 1996 State College, PA (children, wheeze)	1.59 (0.93, 2.70) (1-d)	$PM_{2.1}$ 23.5 (max 85.8)
Neas et al., 1996 State College, PA (children, cold)	1.61 (1.21, 2.17) (0-d)	$PM_{2.1}$ 23.5 (max 85.8)
Schwartz and Neas, 2000 Six Cities reanalysis*** (children, cough)	1.28 (0.98, 1.67)	$PM_{2.5}$ (same as Six Cities)
Schwartz and Neas, 2000 Six Cities reanalysis*** (children, lower respiratory symptoms)	1.61 (1.20, 2.16)	$PM_{2.5}$ (same as Six Cities)

* Studies in italics available in 1996 CD. Studies in bold indicate studies that support choice of proposed locations to include in $PM_{2.5}$ risk analyses based on measure of statistical power of study, where available, and availability of sufficient recent $PM_{2.5}$ air quality data.

** Mean (minimum, maximum) 24-h PM level shown in parentheses unless otherwise noted.

***Six cities studies included the following locations: Boston, Knoxville, Portage, St. Louis, Steubenville, and Topeka.

****The health endpoint in these studies is emergency department visits.

Exhibit C.3. Monitor-Specific Minimum Quarterly Count and Total Annual Count of Days With Measured $PM_{2.5}$ for Areas with $PM_{2.5}$ C-R Functions of Relatively Greater Statistical Power.[1]

Urban Area	Counties	Year	Monitor ID[2]	Min # per quarter	Total Count
Philadelphia	Philadelphia	2000	**421010004881011**	45	266
			421010020881011	10	91
			421010024881011	12	95
			421010027881011	0	47
			421010047881011	7	91
			421010136881011	45	246
Philadelphia Region	Philadelphia, Bucks, Delaware, Montgomery, Camden (NJ), Gloucester (NJ), Burlington (NJ).	2000	**340070003881011**	22	108
			3400700038810124	10	53
			340071007881011	22	103
			340155001881011	25	108
			420170012881011	19	94
			420450002881011	27	112
			4204500028810124	1	6
			420910013881011	20	100
			421010004881011	45	266
			421010020881011	10	91
			421010024881011	12	95
			421010027881011	0	47
			421010047881011	7	91
			421010136881011	45	246
Los Angeles	Los Angeles	2000	**060370002881011**	74	326
			060371002881011	3	70
			060371103881011	59	275
			060371201881011	25	107
			060371301881011	28	116
			060371601881011	27	112
			060372005881011	25	110
			060374002881011	56	259
			0603790028810115	23	107
Phoenix	Maricopa	1999	**040139990881011**	21	101
			040139991881011	43	263
			040139992881011	13	104
			040139997881011	28	258
San Jose	Santa Clara	2000	**060850004881012**	14	179
			0608520038810113	11	188
Boston	Middlesex, Suffolk, Norfolk.	2000	250170008881011	0	34
			250171102881011	11	67
			250210007881011	0	61
			2502500028810113	7	83
			2502500278810113	13	100
			250250027881012	0	31
			250250042881011	56	250
			250250043881011	13	87
Detroit	Wayne	2000	**261630001881011**	81	339
			2616300018810124	12	54
			261630015881011	28	119

Urban Area	Counties	Year	Monitor ID[2]	Min # per quarter	Total Count
			261630016881011	74	324
			261630019881011	0	67
			261630025881011	25	111
			261630033881011	23	107
			261630036881011	16	102
St. Louis	St. Louis, Franklin, Jefferson, St. Charles, Clinton (IL), Madison (IL), Monroe (IL), St. Clair (IL), St. Louis city.	2000	171190023881011[3]	28	115
			171191007881011[3]	27	119
			171192009881011	29	120
			171193007881011[3]	28	117
			171630010881011[3]	27	113
			171634001881011	24	106
			290990012881011	18	103
			291831002881011	27	113
			291892003881011	28	118
			291895001881011	28	118
			295100007881011	6	269
			295100085881011	88	358
			295100086881011	67	322
			295100087881011[3]	66	336
Seattle	King	2000	**530330017881011**	24	108
			530330021881011	82	351
			530330024881011	25	115
			530330027881011	28	114
			530330032881015	0	53
			530330033881011	20	110
			530330037881011	0	16
			530330057881011[3]	81	352
			530330080881011	80	339
			530332004881011	0	9
Portage	Adams, Columbia, Dane, Dodge, Green Lake, Jefferson, Juneau, Marquette, Sauk.	2000	550250025881011	28	115
			550250047881011	28	119
			550270007881011	28	117
			5505500088810113	30	123
		1999	550250025881011	25	109
			550250047881011	23	112
			550270007881011	15	87
			5505500088810113	28	114

Note: Only federal reference method monitors are shown here and are potential candidates for use in the risk analyses.

1. See section 3.2 for a discussion of the criteria used to select potential locations for the risk analyses based on study power. In the absence of sufficient year 2000 PM2.5 data, 1999 PM2.5 data were examined.

2. Monitor sites in bold met the criteria of not being identified as "highest concentration" or "source impact" and of having at least 11 daily values in each quarter. In order for a location to be included in the risk analyses, however, there must be at least one monitor with at least 11 daily values per quarter and 122 daily values per year.

3. These monitors have objectives designated as either "highest concentration" or "source impact." These monitors are excluded from potential use in the risk analyses.

4. These monitors are located at the same site as the monitor directly above it. If two monitors are collecting $PM_{2.5}$ data at the same site, the monitor with more complete data is used.

5. This monitor is located in a separate air basin from the other Los Angeles County monitors and will be excluded from the analysis. Only the portion of Los Angeles County in the South Coast air basin will be examined in this analysis.

Exhibit C.4. Estimated Increased Mortality per Increments in Long-Term Mean Levels of $PM_{2.5}$ from U.S. Studies.

Study Locations and Population*	Change in Health Indicator per 10 $\mu g/m^3$ Increment in $PM_{2.5}$ Relative Risk (95% CI)	Range of City PM Levels Means ($\mu g/m^3$)
Dockery et al. (1993) *Six U.S. Cities* *Age 25-74*	*1.13 (1.04-1.23)*	*11-30*
Krewski et al. (2000) **Six City Reanalysis** **Age 25-74**	1.13 (1.04-1.23)	11-30
Pope et al. (1995) *50 U.S. cities* *Age 30 and older*	*1.07 (1.04-1.10)*	*9-34***
Krewski et al. (2000) **ACS Study Reanalysis** **63 U.S. cities** **Age 30 and older**	1.05 (1.02-1.07)	10-38

* Studies in italics available in 1996 CD. Studies in bold indicate proposed source of relative risk estimates for use in risk analyses.

** Range in PM levels for Pope et al. (1995) is based on annual median, rather than mean.