Responses to Significant Comments on the 1996 Proposed Rule on the National Ambient Air Quality Standards for Particulate Matter

(December 13, 1996; 61 FR 65638)

Docket Number A-95-54

U.S. Environmental Protection Agency
July 1997

Table of Contents

List of Acronyms

The following acronyms have been used for the sake of brevity in this document:

Act	Clean Air Act
BS	British Smoke
CAA	Clean Air Act
CASAC	Clean Air Scientific Advisory Committee
COH	Coefficient of haze
COPD	Chronic Obstructive Pulmonary Disease
EPA	Environmental Protection Agency
FRM	Federal Reference Method
IMPROVE	Interagency Monitoring of Protected Visual Environment
NAAQS	National ambient air quality standards
NAPAP	National Acid Precipitation Assessment Program
OR	Odds Ratio
PM	Particulate matter
RR	Relative Risk
SMSA	Standard Metropolitan Statistical Area

Frequently Cited Documents

The following documents are frequently cited throughout EPA's response to comments, often by means of the short names listed below:

Criteria Document:	U.S. Environmental Protection Agency (1996a). Air Quality Criteria for Particulate Matter. Research Triangle Park, NC: National Center for Environmental Assessment. Office of Research and Development. April 1996. EPA report no. EPA/600/P-95/001aF-cF.
Preamble to the final rule:	Preamble to the Final Rule on the Review of the National Ambient Air Quality Standards for Particulate Matter; to be published in the *Federal Register* on July 18, 1997.
Proposal notice:	National Ambient Air Quality Standards for Particulate Matter: Proposed Decision. 61 FR 65638, December 13, 1996.
Staff Paper:	U.S. Environmental Protection Agency (1996b). Review of the National Ambient Air Quality Standards for Particulate Matter: Assessment of Scientific and Technical Information. OAQPS Staff Paper. Research Triangle Park, NC: Office of Air Quality Planning and Standards. July 1996. EPA report no. EPA-452/R-96-013.
Summary of Comments:	Abt Associates (1997). Final Summary of Significant Written Comments on the 1996 Particulate Matter NAAQS Proposal - Issue Report. EPA Docket # A-95-54. TRC Environmental Corporation (1997). PM Monitoring Technical Support Document (Part I). Summary of Public Comments on Proposed PM Monitoring Regulations (40 CRF Part 50 Appendix L, 40 CFD Part 53, 40 CFR Part 58).

Responses to Significant Comments on the 1996 Proposed Rule on the National Ambient Air Quality Standards for Particulate Matter

I. INTRODUCTION

This document, together with the preamble to the final rule on the review of the national ambient air quality standards (NAAQS) for particulate matter (PM) and several separate documents referred to below, presents the responses of the Environmental Protection Agency (EPA) to the more than 50,000 public comments received on the 1996 PM NAAQS proposal notice. All significant issues raised in the public comments have been addressed.

As reflected in the table of contents for this document, responses are organized by topics, which correspond to specific sections of a companion document that has been placed in the docket, the Summary of Significant Comments on the 1996 Particulate Matter NAAQS Proposal -- Organized by Issue (henceforth the "Summary of Comments").[1] Due to the large number of comments that addressed similar issues, as well as the sheer volume of the comments received, this response-to-comments document does not generally cross-reference each response to the commenters who raised the particular issue involved, although commenters are identified in some cases where they provided particularly detailed comments that were used to frame the overall response on an issue.

This document refers as appropriate to various support documents, available in the docket, that have been prepared to assist in presenting the more technical aspects of the Agency's responses. A complete list of references, including these support documents, is presented at the end of this document. In addition, separate summary and response to public comments documents have been prepared for issues raised in comments on the proposed new reference method for measuring fine particles as $PM_{2.5}$.[2]

The responses presented in this document, including its appendices, and in the separate documents referred to above are intended to augment the often extensive responses to comments that appear in the preamble to the final rule or to address comments not discussed in the preamble. Although portions of the preamble are paraphrased in this and other documents where

---

[1]A second summary document, the Summary of Significant Comments on the 1996 Particulate Matter NAAQS Proposal - Organized by Commenter, has also been placed in the docket to facilitate the review of comment summaries by commenter as well as by issue.

[2]The latter documents also includes responses to public comments on related proposals to revise 40 CFR Parts 53 and 58.

useful to add clarity to responses, the preamble itself remains the definitive statement of the basic rationale for the revisions to the standards adopted in the final rule.

In many instances, particular responses presented in the above documents include cross references to responses on related issues, either in those documents or in the preamble to the final rule.[3] In view of the large number of comments received, the cross references may not always reflect the extent to which information relevant to a particular comment is contained in responses to other comments. Accordingly, the above documents as a group, together with the preamble to the final rule, should be considered collectively as EPA's response to all of the comments submitted.

II. RESPONSES TO SIGNIFICANT COMMENTS ON PROPOSED PM STANDARDS

A. Primary PM Standards

1. General comments on proposed primary standards

A large number of comments on the proposed primary standards for PM were very general in nature, basically expressing one of two substantively different views: (1) support for revisions to the primary standards by addition of new standards for $PM_{2.5}$ (either the proposed standards or alternative standards more or less stringent than those proposed); or (2) opposition to the addition of $PM_{2.5}$ standards, with or without any revisions to the current PM_{10} standards. Many of these commenters simply expressed their views without stating any rationale, while others gave general reasons for their views but without reference to the factual evidence or rationale presented in the proposal notice as a basis for the Agency's proposed decision. Comments of this nature on the PM proposal notice are summarized in section II.A.1 of the Summary of Comments, and those addressing both the PM and O_3 proposal notices are summarized in section I.A.

The preamble to the final rule in its entirety presents the Agency's response to these general views. More specifically, section II of the preamble responds to views that are health-based, including those related to the following factors: 1) the strengths and limitations of the scientific evidence on the effects of PM; 2) the need for and appropriateness of revising the standards by adding primary standards for $PM_{2.5}$ now, as opposed to waiting for additional research and monitoring for $PM_{2.5}$; and 3) the advice of the Clean Air Scientific Advisory Committee (CASAC) on the adequacy of the scientific evidence available for making a decision on the standards and individual CASAC Panel members' personal views on the standards.

---

[3]The terminology used in the preamble to the final rule as it appears in the Federal Register refers to various named sections of the preamble as "units." This response to comments document refers to these units as "sections" of the preamble.

Sections IV and VIII of the preamble respond to comments on certain legal and procedural issues.

2. Specific comments on proposed primary $PM_{2.5}$ standards

A large number of comments addressed specific elements of the proposed primary $PM_{2.5}$ standards, including the averaging time, level, and form of the standards, and the provisions relating to interpretation of the standards as specified in a revised Appendix K in the proposal and in Appendix N in the final rule. These comments are generally summarized in sections II.A.2.a and b of the Summary of Comments, and responses to the key issues raised in these comments are presented in sections II.C,D,E, and F and section V of the preamble to the final rule. More specific responses to the full range of significant issues raised in these comments are presented below.

a. Indicator for fine particles

A broad range of public comments were received in this area. A number of commenters raised various issues with regard to EPA's proposed general mass indicator for fine particles, $PM_{2.5}$. Some commenters expressed the view that no fine particle mass indicator is warranted for various reasons. Others provided comments as to how EPA should modify the proposed $PM_{2.5}$ indicator [and associated Federal Reference Method (FRM) monitor] to address various concerns. EPA notes that it considered a variety of indicators (e.g., PM_{10}, $PM_{2.5}$, PM_1, chemical component) during the review, as discussed in the Staff Paper. The comments are summarized primarily in section II.A.2.a.(1) of the Summary of Comments and significant comments are highlighted in the preamble to the final rule. In addition to the responses contained in sections II-B and II-C-1 of the preamble to the final rule, EPA provides the following responses to specific issues.

i. <u>Use of fine particle mass indicator</u>

The following comments raise issues as to why a fine particle mass indicator is not warranted:

(1) *Comment:* Many commenters expressed the view that a fine particle mass indicator in general, or a $PM_{2.5}$ indicator in particular, is not supported by the available scientific evidence. These views are based on assertions that there are too few $PM_{2.5}$ health effects studies and/or that the available $PM_{2.5}$ studies are too uncertain or flawed (e.g., confounded by other pollutants and/or weather, biased by measurement error, inadequate to prove causality) to be used as a basis for setting fine particle standards.

Response: See sections II.B and II.C.2 of the preamble to the final rule, for a general response to the overall adequacy of the scientific data base with regard to the need to revise the PM standards, including a discussion of general and specific issues with respect to the available epidemiological information on the effects of PM.

As outlined in sections II.B and II.C of the preamble to the final rule, the Staff Paper concludes that continued use of PM_{10} as the *sole* indicator for the PM standards would not provide the most effective and efficient protection from the health effects of PM (U.S. EPA, 1996b, pp. VII-4-11). Based on the recent health effects evidence and the fundamental physical and chemical differences between fine and coarse fraction particles, the Criteria Document and Staff Paper conclude that fine and coarse fractions of PM_{10} should be considered separately (U.S. EPA, 1996a, p. 13-93; 1996b, p. VII-18). Taking into account such information, CASAC found sufficient scientific and technical bases to support establishment of separate standards relating to these two fractions of PM_{10}. Specifically, CASAC advised the Administrator that "there is a consensus that retaining an annual PM_{10} NAAQS . . . is reasonable at this time" and that there is "also a consensus that a new $PM_{2.5}$ NAAQS be established" (Wolff, 1996b).

Beyond the general points about the basis for any revisions discussed in section II.A.1 above, these commenters argued either that the available epidemiological data did not provide a basis for separating fine and coarse fraction particles, or that there were not enough fine particle studies to support selecting standard levels. Most of these commenters also expressed concerns that there were insufficient ambient fine particle data by which to evaluate the relative protection afforded by new standards.

EPA notes that issues relating to the basis for separating PM_{10} fractions were addressed in the Criteria Document and/or Staff Paper assessments, and these perspectives were also available for CASAC consideration in developing its recommendations. As noted in the Staff Paper, the main basis for separating the fine and coarse fractions of PM_{10} is that, because they are fundamentally different PM components with significantly different physico-chemical properties and origins (U.S. EPA 1996b, Section V.D), separate standards would permit more effective and efficient regulation of PM. While the difficulty of separating these classes of particles in the epidemiological studies is recognized in the Staff Paper assessment, the preponderance of the available evidence suggests that strategies to control fine particles will more effectively reduce population exposure to substances associated with health effects in the recent epidemiological studies. Although the number of studies using fine PM indicators is more limited than for PM_{10}, there are more than 20 community studies showing significant associations for a consistent set of mortality and morbidity effects. A substantial subset of these studies (Tables V-12 to V-13; U.S. EPA, 1996b) provides a sufficient quantitative basis for selecting standard levels, without the need to rely on estimates based on $PM_{2.5}/PM_{10}$ ratios.

See also section II.A.4 and Appendices A, B and D below on interpretation of the epidemiological studies and issues such as consistency, coherence, confounding and measurement error.

(2) *Comment:* Many commenters expressed the view that an indicator based on fine particle mass is inappropriate because it does not differentiate specific causal agents within the

mix of fine particle components. A number of commenters expressed concerns that various portions of fine particles might not be responsible for any observed effects. As a result, local and/or regional control strategies may be misdirected toward fine particle components that are not related to health effects. In addition, a commenter (API, IV-D-2250) asserted that the proposal notice understates the variety among the $PM_{2.5}$ constituents.

Response: See preamble to the final rule, sections II.B and II.C.2.

Consistent with CASAC advice, this review considered the merits of alternative indicators, including $PM_{2.5}$ mass as well as indicators based on specific components of PM. The Criteria Document extensively evaluated health effects information on many specific components, including sulfates, acids, nitrates, organics, and transition metals. In addition, during the review EPA specifically highlighted the diversity within the fine and coarse particle mix (e.g., see Table IV-2 in the Staff Paper). Based on this extensive review, staff concluded, with CASAC concurrence, that a $PM_{2.5}$ indicator was appropriate. In so doing, staff also specifically considered the likely effect of the use of such an indicator for control of PM components and key gas-phase precursors of PM.

As noted in the preamble to the final rule, EPA continues to conclude that it is appropriate to control fine particles as a group, as opposed to singling out particular components or classes of fine particles. The more qualitative scientific literature, evaluated in Chapter 11 of the Criteria Document and summarized in Section V.C of the Staff Paper, has reported various health effects associated with high concentrations of a number of fine particle components (e.g., sulfates, nitrates, organics, transition metals), alone or in some cases in combination with gases. Community epidemiological studies have found significant associations between fine particles and/or PM_{10} and health effects in various areas across the U.S. where such fine particle components correlate significantly with particle mass. As noted in the Staff Paper, it is not currently possible to rule out any one of these components as contributing to fine particle effects. Thus, the Administrator finds that the present data more readily support a standard based on the total mass of fine particles.

Information suggesting that observed PM-related health effects are related to specific components of regional or local concern is not nearly as extensive as information relating health effects to particle mass (U.S. EPA 1996a, chapters 12 and 13). In addition, EPA is required to establish national standards, and states and local governments are not precluded from establishing additional standards that may be deemed appropriate to address concerns about particular classes of fine particles. In specifying a precise size range for a fine particle standard, both the staff and CASAC recommended $PM_{2.5}$ mass concentration as the indicator of fine particles (Wolff, 1996b). As noted in the Staff Paper, $PM_{2.5}$ encompasses all of the potential agents of concern in the fine fraction, including most of the particle number in the entire PM distribution as

well as most sulfates, acids, fine particle metals, organics, ultrafine particles and most of the aggregate surface area (U.S. EPA 1996b, p.VII-15).

The central question of which particulate components to regulate has been an issue since the inception of the first PM standards. Other ambient pollutants (e.g., NO_2 or CO) are uniquely defined as individual chemicals, whether or not they serve as proxies for a larger class of substances (e.g., ozone as an index of photochemical oxidants). Regulating PM in general, as opposed to multiple chemical components of PM, raises questions as to whether particulate components of varying composition, size, and other physicochemical properties are likely to produce identical effects.

Both EPA's past and present regulatory experience with PM control programs and its successive reviews of the standards have reaffirmed the wisdom of retaining standards that control particles as a group, rather than eliminating such standards and waiting for scientific research to develop information needed to identify more precise limits for the literally thousands of particle components that may potentially be of concern. Each such decision recognized the possibility that potentially less harmful particles might be included in the mix that was regulated, but concluded that the need to provide protection against serious health effects nonetheless required action under section 109 of the Act. The success of this approach is evident in early U.S. control programs that dramatically reduced "smoke" and "TSP" in major cities in the 1960's and 1970's and in the continued improvement in air quality through the current PM standards. The major refinements that have been recommended through the course of reviews of PM standards have been to improve the focus of control efforts by defining scientifically based size classes (i.e., moving from TSP to PM_{10} and now, $PM_{2.5}$) that will permit more effective and efficient regulation of those fractions most likely to present significant risks to health and the environment.

As discussed in the preamble to the final rule, the current review has examined the available evidence to determine whether it would tend to support inclusion or exclusion of any physical or chemical classes of PM, for example sulfates, nitrates, or ultra-fine particles. That examination concludes that, while both fine and coarse particles can produce health effects, the fine fraction appears to contain more of the reactive substances potentially linked to the kinds of effects observed in the recent epidemiological studies (U.S. EPA 1996b, Section V.F.). However, the available scientific information does not rule out any one of these components as contributing to fine particle effects. Indeed, it is reasonable to anticipate that no single component will prove to be responsible for all of the effects of PM.

EPA recognizes that whether the standards are set for PM_{10} only or also for fine particles, there are uncertainties with respect to the relative risk presented by various components of PM. In this regard, the EPA is placing greater weight on the concern that by failing to act now, the PM NAAQS would not control adequately those components of air pollution

that are most responsible for serious effects, than on the possibility they might also control some component that is not.

EPA also recognizes that different components may be of interest in different areas (e.g., woodsmoke, sulfates, etc). EPA notes that the epidemiological studies were conducted in locations around the nation with different mixes of components but with reasonably similar results in terms of responses. Consequently, EPA selected a mass-based fine particle indicator. In attaining the $PM_{2.5}$ standards, areas may choose to focus their control strategies on these different components. However, EPA believes that the general particle indicator remains an appropriate approach to protecting public health.

(3) *Comment*: EPA should use PM_{10} as the indicator to control fine particle mass since PM_{10} correlates as well with reported health effects as do various fine particle indicators, including $PM_{2.5}$, and /or because current PM_{10} control programs already function in some areas to control fine particles.

Response: See preamble to the final rule, sections II.B and II.C.

Based on both the staff review (U.S. EPA, 1996b, p. VII-3) and the recommendations of some commenters (e.g., California EPA, IV-D-2251), EPA has considered two alternative approaches for providing additional health protection in revising the standards: 1) adopt more protective PM_{10} standards and 2) develop separate standards for the major components of PM_{10}, including $PM_{2.5}$. Conceptually, the first approach would give weight to the view that standards should be based on pollutant indicators for which the most data have been collected, with less consideration of the evidence that suggests that the current standards provide adequate protection against the effects of coarse particles, and that tightening the current standards to control fine particles would place unnecessary requirements on coarse particles. Because the PM_{10} network is in place, a more health protective PM_{10} standard would also respond to commenters who have expressed a desire for more immediate implementation of revised standards. The second approach is based on the view that in the long run, more effective and efficient protection can be provided by separately targeting appropriate levels of controls to fine and coarse PM.

The Staff Paper examined this issue in detail (U.S. EPA 1996b, p. VII-3 to 11), and concluded that the available information was sufficient to develop separate indicators for fine and coarse fractions of PM_{10}, based on the recent health evidence, the fundamental differences between fine and coarse fraction particles, and implementation experience with PM_{10}. Further, the staff concluded that:

> [C]onsideration of comparisons between fine and coarse fractions suggests that fine fraction particles are a better surrogate for those particle components linked to mortality and morbidity effects at levels below the current standards. In contrast, coarse fraction particles are more likely

> linked with certain effects at levels above those allowed by the current PM_{10} standards. In examining alternative approaches to increasing the protection afforded by PM_{10} standards, the staff concludes that reducing the levels of the current PM_{10} standards would not provide the most effective and efficient protection from these health effects. [U.S. EPA 1996b; p. 7-45.]

As discussed in section II.C of the preamble to the final rule, EPA concludes that it is more appropriate to provide additional protection against the risk posed by PM by adding new standards for the fine fraction of PM_{10}, than by tightening the current PM_{10} standards. Although fewer epidemiological studies have used $PM_{2.5}$ and other fine particle indicators (e.g., sulfates, acids), there are nonetheless significant indications from the scientific evidence drawn from the physicochemical studies of PM, air quality and exposure information, toxicological studies, and respiratory tract deposition data that this approach will provide the most effective and efficient protection of public health.

(4) *Comment:* Some commenters argued that the results of a study by Schwartz et al. (1994), which EPA used in support of a $PM_{2.5}$ standard, more readily suggest that PM_{10} is a better indicator of PM in health effects studies.

Response: Schwartz et al. (1994) analyzed symptom diary data from children in six cities for associations with daily measurements of ambient air pollutants. In this report, results were presented for associations between three symptom groups and air pollutants. The findings for PM and PM components are presented in the table below.

SUMMARY OF RESULTS FROM SIX CITIES SYMPTOMS STUDY
(Schwartz et al., 1994)

	Cough OR (95% CI)	**Lower respiratory symptoms** OR (95% CI)	**Upper respiratory symptoms** OR (95% CI)
PM_{10} (30 μg/m^3)	1.27 (1.06-1.52)	1.53 (1.20-1.95)	1.22 (0.98-1.52)
$PM_{2.5}$ (20 μg/m^3)	1.19 (1.01-1.42)	1.44 (1.15-1.82)	1.22 (1.00-1.49)
Fine particle sulfur (5 μg/m^3)	1.23 (0.95-1.59)	1.82 (1.28-2.59)	reported not significant
Nephelometry (1 km^{-1})	1.21 (1.02-1.45)	1.36 (1.14-1.63)	reported not significant
Aerosol acidity (25 nmol/m^3)	1.06 (0.87-1.29)	1.05 (0.25-1.30)	1.06 (0.98-1.15)

Commenters have focused on the following conclusion reached by the authors based on these results: "There was no evidence that other measures of particulate pollution including aerosol acidity were preferable to PM_{10} in predicting incidence of respiratory symptoms." (Schwartz et al., 1994) However, EPA notes the same could as easily be said about $PM_{2.5}$. The Staff Paper and Criteria Document note the difficulty in separating the effects associated with these overlapping indicators of PM, particularly in comparing fine particle components such as sulfates or acids with $PM_{2.5}$, and $PM_{2.5}$ with PM_{10}. In this study, $PM_{2.5}$ apparently comprises about 2/3 of PM_{10} mass, which suggests that any attempt to separate the two in this study is questionable, despite the authors' conclusions and commenters' emphasis on this portion of the results. (These commenters in general do not note that the PM_{10} concentrations associated with effects in this study are generally well below those permitted by the current standards.)

In evaluating the usefulness of $PM_{2.5}$ as an indicator, the Staff Paper focused not on comparisons between PM_{10} and $PM_{2.5}$, but on comparing the relative effects associations between fine and coarse fraction particles that, taken together, comprise PM_{10}. Although the Schwartz et al. (1994) work did not report results separately for $PM_{2.5}$ and $PM_{10-2.5}$, a number of other studies, including the six city mortality study, did provide useful results in this regard (U.S. EPA, 1996b; Section V.D). In essence, the staff assessment found that fine particles or fine particle components are generally stronger predictors of adverse health effects where such comparisons are possible. As summarized in the Staff Paper (pp. V-58 to V-67), EPA believes that the decision to establish a $PM_{2.5}$ standard is well supported by the available science.

ii. Inclusion of Constituents in Measured $PM_{2.5}$

A number of commenters raised made alternative suggestions for how EPA should modify the proposed $PM_{2.5}$ indicator (and associated FRM monitor) to better address the collection of fine particle components of concern, with some recommending elimination of one or more components, and others recommending more complete capture of nitrates and certain other semi-volatile components in the FRM. Specific comments in this area are summarized and responded to below. As a general response, however, EPA notes that the fundamental approach to selecting a fine particle indicator was based on consistency with the monitoring used in the underlying epidemiological studies. The specific philosophy and the approach used for the FRM was reviewed by a CASAC Technical Subcommittee for Fine Particle Monitoring at a public meeting held March 1, 1996, in Chapel Hill, NC. This CASAC subcommittee consisted of monitoring and other experts who concurred with EPA's proposed approach, recognizing the inherent challenges of being consistent with the studies on which the levels of the standards are based and the goal of fully characterizing the aerosol in many diverse settings across the country using a consistent approach required by an equitable national standard (Price, 1996). The subcommittee found that, "Under the circumstances, EPA has made an appropriate choice to establish a good practice standard for filter sampling and analysis technology" (Price 1996). Furthermore, the subcommittee endorsed EPA's specific approach with respect to nitrates: "Since the recent epidemiological studies have used a variety of methods with different

performance characteristics, no one FRM can match them all; however, [EPA's FRM approach] matches most in the choice not to use a more complex design that includes denuding and backup filtration to improve the sampling of the ambient particulate nitrate compounds" (Price 1996, emphasis added).

(1) *Comment*: If a fine particle standard is promulgated, the EPA should exclude nitrates in the definition of $PM_{2.5}$, or EPA should allow States to set standards for PM excluding nitrates, due to the lack of toxic effects from airborne nitrates. One commenter (Ohio Edison, IV-D-2275) asserts that failure to exclude nitrates from the definition of particulate matter would lead to conflict with a previous EPA decision to not revise the current national ambient air quality standards for NO_2. Another commenter (PG&E, IV-D-2183) recommends that nitrates should be excluded from fine PM mass collected on the basis of its assessment of a list of studies from available effects literature on particulate and gas phase inorganic nitrates.

Response: As discussed in the preamble to the final rule, EPA continues to conclude that it is appropriate to control fine particles as a group, as opposed to singling out particular components or classes of fine particles. The more qualitative scientific literature, evaluated in Chapter 11 of the Criteria Document and summarized in Section V.C of the Staff Paper, has reported various health effects associated with high concentrations of a number of fine particle components (e.g., sulfates, nitrates, organics, transition metals), alone or in some cases in combination with gases. Community epidemiological studies have found significant associations between fine particles or PM_{10} and health effects in various areas across the U.S. where such fine particle components correlate significantly with particle mass. As noted in section II.C of the preamble, it is not possible to rule out any one of these components as contributing to fine particle effects.

With respect to the specific comment that the inclusion of nitrates would be at odds with other NAAQS decisions, it is important to distinguish nitrate-bearing particles from NO_2, which is a gas, as well as from nitric acid when it occurs as a vapor. While commenters are correct that particulate nitrates and nitric acid are generally formed from atmospheric conversion of nitrogen oxides emissions, (just as sulfates are formed from atmospheric transformations of sulfur oxides emissions), the recent reaffirmation of the NO_2 standards was based principally on the effects of NO_2 itself, and not on the atmospheric transformation products of NO_2. NO_x is already recognized as a precursor to ozone and PM_{10}. Therefore, EPA does not consider the existence of a separate gaseous NO_2 standard to be relevant to the inclusion of nitrate particles in the indicator for PM. Whether or not there is a correlation between atmospheric levels of NO_2 and nitrate particles, EPA believes that it is preferable to rely on health effects studies that used exposures to nitrate- or nitric acid-bearing particles in order to draw any conclusions regarding the health effects of particles with associated nitrates or nitric acid. Thus, EPA does not agree with the premise that the studies of NO_2 exposure are directly relevant to a discussion of toxic effects from exposure to nitrates or nitric acid. EPA sees no cause for concern about a regulatory conflict between the NAAQS for PM and NO_2. The primary

PM NAAQS is based on studies of the health effects from exposure to PM, including acid aerosols. This decision is independent of any decisions made regarding the NO_2 NAAQS.

With regard to the claim that EPA did not consider certain studies, EPA disagrees. Chapter 11 of the PM Criteria Document, which discusses controlled human and toxicological studies of PM components, specifically cross references documents in which the specific studies on the effects of nitrates have been reviewed, including the 1982 PM Criteria Document, the 1989 EPA Acid Aerosols Issues Paper, and the 1993 NO_x Criteria Document (U.S. EPA, 1996a: p 11-5). The document notes that the more limited recent information on nitrates is summarized in sections on multi-component studies (such as that of Kleinman et al., 1995). The EPA staff considered the results of relevant studies on nitrates in reaching conclusions on the appropriate indicator for fine particles. Indeed, the studies examined by EPA also included nitrate studies that commenters chose not to cite.

In examining the list of studies submitted for further consideration by PG&E (IV-D-2183), EPA notes that two of the nine (Braun-Fahrlander et al., 1992, and Dockery et al., 1992) did not use separate quantitative measurements of nitrates or nitric acid aerosols, but rather relied on measurements of NO_2. EPA did consider these studies with regard to PM, but they are not considered to be directly relevant to a discussion of health effects from nitrate-bearing particles. However, EPA notes that in the Braun-Farhlander et al. (1992) study the annual average of NO_2 was associated with the duration of upper respiratory symptoms.

Because relatively few epidemiological studies have included separate quantitative measures of nitrate particles or nitric acids, the Criteria Document did not set apart a discussion of nitrates or nitric acids from the overall review of health effects from acid aerosol exposures. The commenter cites three epidemiologic studies (Ostro et al., 1991; Burnett et al., 1989; Burnett et al., 1994) as finding no associations between nitrate levels and asthma symptoms or lung function. EPA notes that commenters omitted other studies cited in the Criteria Document that reported positive associations with nitrates. Fine particulate nitrates were associated with increased symptoms and bronchodilator use in Perry et al. (1982), and associations were found with aerosol nitrate and nitrous acid by Hoek and Brunekreef (1994). In addition, several epidemiological studies reporting statistically significant associations between PM exposures and serious health effects (e.g., mortality) were conducted in areas in which nitrate levels are expected to be relatively high (e.g., Los Angeles, Utah Valley). Nitrate particles are expected to be present to some degree in most urban areas in the U.S. (See Criteria Document Chapter 6 and Figures 6-85a-c and Staff Paper Figure IV-3). Although nitrates were not measured quantitatively, these epidemiological studies lend support to the inclusion of the fraction of nitrate present in these types of studies. However, EPA emphasizes that, in reviewing the available epidemiological data, no evidence has been found to give the agency cause to distinguish this subset of particulate matter for the purposes of regulation.

In addition to the epidemiological studies, the commenter (PG&E) also included several toxicological and clinical studies of nitrate or nitric acid exposure. The toxicological studies were available for previous reviews of air quality standards. In the PM Criteria Document, almost all of the more recent clinical or toxicological studies on acid aerosols that were available for review used sulfuric acids or sulfate particles. It is of note that a pattern of positive and negative responses to nitrates mirrors the much more numerous findings for sulfates. Human studies of nitric acid vapor exposure have generally not found significant effects on spirometric measures of lung function. The PM Criteria Document cites the toxicological studies listed by the commenter and recognizes that health effects were found in studies that used high nitrate concentrations (p. 11-7), as observed by the commenter. Commenters are incorrect, however, in stating that nitrates have been shown to be uniformly less toxic than other PM components. Although limited studies in humans (cited in both the Acid Aerosol Issue Paper and the NOx Criteria Document) show minimal effects on spirometric measures of lung function and respiratory symptoms, Kleinman et al. (1995) observed an increase in lung permeability in relation to exposure to a high (350 $\mu g/m^3$) concentration of particulate nitrate. These observations indicate that nitrate cannot be ignored and treated as an "inert" particle.

Commenters also submitted two unpublished abstracts of studies that were not available for inclusion in the Criteria Document. Disregarding the appropriateness of relying on such more recent studies, the results of Balmes et al., who found no significant incremental effects in particular endpoints after short-duration exposures of humans to nitric acid vapor and nitrates, are consistent with previous controlled human exposures to these materials. However, as is the case for similar findings for acid sulfates, such results cannot be said to disprove any role of nitrates in the observed findings of epidemiologic studies, which include different endpoints, far larger numbers of people, significantly more sensitive populations, and considerably longer exposure durations than can be achieved in controlled human studies. The unpublished results by Kleinman et al. in animals apparently find effects of nitrates at levels as low as 100 $\mu g/m^3$ and increased potency with respect to other PM components at levels of 350 $\mu g/m^3$. If the results of this study were to be considered, assuming they were sustained following peer review, publication, and inclusion in the next criteria review, they would serve to add markedly to, not diminish, concerns about nitrates. In addition, commenters did not cite another recent study of nitric acid vapor that was published after the NO_x Criteria Document was completed (Schlesenger et al., 1994). This study found that inhalation of nitric acid vapor levels as low as 50 $\mu g/m^3$ may adversely affect pulmonary health in animals by altering the production of superoxides and the release of tumor necrosis factor by alveolar macrophages. In essence, the mixed results of these recent findings, if fully considered, would not support commenters' recommendation to exclude nitrates.

Based on an examination of the information submitted by the commenter, as well as the earlier staff integrated assessment of the available health effects information (largely the same information), EPA maintains its conclusion that the available evidence is not

sufficient to exclude nitrates or any other class of fine particles that are collected by PM monitors comparable to those used in the recent epidemiological studies.

(2) *Comment:* A number of commenters expressed the view that semivolatile components of fine particles, such as nitrates and organics, should be more completely captured than is done in the proposed Federal Reference Method Sampler. Other commenters suggested that water should be excluded.

Response: As noted above, EPA believes the available evidence supports the inclusion of nitrates and other semivolatile species, to the extent they were collected in the underlying community epidemiological studies. EPA has developed its monitoring approach with this objective in mind. While acknowledging the FRM may involve some loss of semivolatile substances, EPA does not believe it would be appropriate to collect amounts of such materials in significantly greater proportions than did the study monitors. Specifically, as discussed above, the CASAC subcommittee acknowledged the complex technical issue and endorsed EPA's approach with respect to nitrates since it "matches most in the choice not to use a more complex design that includes denuding and backup filtration to improve the sampling of the ambient particulate nitrate compounds" (Price 1996, emphasis added).

In implementing the fine particulate indicator through a FRM, EPA notes that some portion of semivolatile species is included; however, this may vary depending on local conditions (e.g., ambient temperature changes during measurement period, general atmospheric chemistry). EPA does note that the particle mix in certain areas (e.g., Western locations) will probably contain more semivolatile species such as nitrates and organics than others (See Staff Paper Figure IV-3).

In developing the FRM, EPA worked to minimize semivolatile fraction losses and rejected some existing candidate FRM technologies that offered other performance advantages because of semivolatile species losses that are inconsistent with the underlying health database. EPA also notes that other methods that would retain more of the semivolatile material could also capture additional water, which is undesirable for compliance with the primary standard. With respect to addressing this issue, Appendix L specifies that handling of samples requires equilibration under defined conditions to standardize water content in the measurements.

EPA encourages the measurement and speciation of all particulate matter, including the semivolatile fraction, both for defining air quality and for conducting research relevant to subsequent reviews. EPA also notes that full characterization of the chemistry of the aerosol, including semivolatile species and water, is an important component of assessing impact on visibility in the context of Regional Haze programs. Visibility programs have a long-standing monitoring protocol (i.e., the IMPROVE network, which has been in operation since 1987). See responses in Section II.B of this document for additional information.

(3) *Comment:* Some commenters suggested that intrusion of any coarse particles derived from naturally occurring crustal materials should be excluded because, according to the commenters, these particles are not of health concern. Some of these commenters supported exclusion of coarse particles from either PM_{10} or $PM_{2.5}$ standards. Some of these commenters argue that PM_1 (particles with sizes less than a nominal 1 μm) would be a more appropriate indicator for fine particles. Some commenters expressed the view that, if EPA adopts $PM_{2.5}$ standards, the Agency should provide a method that would result in better separation of fine and coarse fraction particles because there may be some intrusion of coarse material into the $PM_{2.5}$ measurement.

Response: See preamble to the final rule, section II.B, for further discussion of relevant health effects information.

EPA disagrees with the comment that crustal particles or "naturally occurring" particles are not of health concern. The preamble to the final rule and the Staff Paper conclude that coarse fraction particles are clearly linked with certain morbidity effects, and CASAC clearly supported retention of a PM_{10} indicator to protect against the effects of coarse fraction particles. The Criteria Document and Staff Paper conclude that epidemiological information, together with dosimetry and toxicological information, support the need for a particle indicator that addresses the health effects associated with coarse fraction particles within PM_{10} (i.e., $PM_{10-2.5}$). As noted above, coarse fraction particles can deposit in those sensitive regions of the lung of most concern. Although the role of coarse fraction particles in much of the recent epidemiological results is unclear, limited evidence from studies where coarse fraction particles are the dominant fraction of PM_{10} suggests that significant short-term exposure effects related to coarse fraction particles include aggravation of asthma and increased upper respiratory illness. In addition, qualitative evidence suggests potential chronic effects associated with long-term exposure to high concentrations of coarse fraction particles.

EPA agrees that it is appropriate to separate fine and coarse fraction particles for regulatory purposes. As stated in the preamble to the final rule, EPA adopted the 2.5 μm limit based on the potential for growth of true fine mode particles into that size, the comparability with epidemiological studies and other monitoring, and the recommendations of CASAC. In the Staff Paper, EPA notes that the $PM_{2.5}$ measurement does have some potential for intrusion of the "tail" of the coarse mode (U.S. EPA 1996b, P. VII-16 and Appendix A). Following the recommendations of CASAC, EPA has selected a FRM with a sharp cut to minimize this potential intrusion of coarse mode particles. While EPA notes that it does not anticipate such intrusions to be significant in most situations, to the extent that problems in this regard occur in some locations, this issue can and should be addressed on a case-by-case basis in the monitoring and implementation programs (e.g., through a policy similar to the natural events policy). In such situations, the programs will be guided by the fact that the PM_{10} standards are intended to protect against the effects of coarse fraction PM, and that fine standards are

directed at smaller sizes. Therefore, consideration of the kinds of exclusions recommended by commenters will be taken up in developing implementation guidance.

EPA notes that a $PM_{2.5}$ indicator is more appropriate at this time than a PM_1 indicator. PM_1 has not been used directly in health studies or widely used in the field, although in most cases mass should be similar for cutpoints of 2.1 or 2.5 μm. While PM_1 could reduce intrusion of fugitive dust, it might also omit portions of hygroscopic acid sulfates in high humidity episodes. In the Staff Paper, EPA notes that of some concern is the theoretical possibility that different flow velocities for the smaller cut might increase the loss of semivolatile materials relative to a larger cut (U.S. EPA 1996b, p. VII-16 and Appendix A).

iii. Consideration of alternative fine particle indicators

(1) *Comment*: One commenter (CAAP, IV-D-8258) expressed the view that EPA should use a particle number indicator rather than particle mass.

Response: Information suggesting that observed PM-related health effects are related to particle number is not nearly as extensive as information relating health effects to particle mass (U.S. EPA 1996a, chapters 12 and 13). Community epidemiological studies have found significant associations between fine particle or PM_{10} mass and health effects in various areas across the U.S. Consistent with the recommendation of 19 of 21 CASAC panel members, EPA proposed to use a $PM_{2.5}$ mass concentration indicator. As noted in the Staff Paper, $PM_{2.5}$ encompasses all of the potential agents of concern in the fine fraction, including most of the particle number in the entire PM distribution as well as most sulfates, acids, fine particle metals, organics, ultrafine particles and most of the aggregate surface area (U.S. EPA 1996b, p.VII-15).

The available information shows that particle number is dominated by directly emitted "ultrafine" (<0.1 μm) particles, which quickly aggregate into larger sizes, as well as particles that form in the air from reaction of gases such as sulfur dioxide. A standard based on numbers of particles would essentially focus exclusively on these ultrafine particles. Both the Criteria Document and the Staff Paper examine the potential contribution of directly emitted ultrafine particles to the observed effects of particulate matter. The Criteria Document points out that such ultrafine aerosols (<0.1 μm) are a class of fine particles that have the potential to cause toxic injury to the respiratory tract as seen in several animal studies (p. 13-76). The Staff Paper assessment includes the following evaluation of potential risk:

> Because of their short lifetime, it is unclear that unaggregated ultrafine particles make up any significant fraction of the mass of fine particles or of PM_{10}, other than in the vicinity of significant sources of ultrafine particles. The relationship between ultrafine numbers (or mass) and the mass of fine or thoracic [PM_{10}] particles found in typical community air

> pollution has not been established. Although the Criteria Document provides little direct information, it might be expected that penetration and persistence of unaggregated ultrafine particles to indoor environments would be limited. For these reasons, it is questionable whether ultrafine aerosols could be playing a major role in the reported epidemiologic associations between the measured mass of fine or PM_{10} particles and health effects in sensitive populations [Staff Paper, p. V-72-73.]

In summary, given their much longer atmospheric lifetime and broader dispersion from source regions, the larger fine particles appear to be of greater risk to public health. Because of the potential toxicity of ultrafine particles and the opportunity for exposure near combustion sources, however, they represent an area where additional research is necessary. In any event, strategies that control fine particles will focus new attention on both directly emitted and atmospherically formed ultrafine particles. EPA believes the available information clearly supports selection of a mass-based indicator, as opposed to a number-based particle standard.

(2) *Comment*: One commenter (U. Of Rochester, IV-D-894) expressed concern that a reduction in accumulation mode mass would lead to an increase in the number of ultrafine particles (aerodynamic diameter <0.1 μm). This would be undesirable if the number of particles, rather than the mass of particles, were the crucial factor in causing health effects.

Response: The suggestion that portions of fine particle mass are a "sink" for ultrafine particles is theoretically correct, but EPA's examination of the issue, which is discussed below, suggests that the practical implications may be quite limited. In essence, while there are situations in which a reduction in "accumulation mode"[4] mass could lead to an increased persistence of directly emitted ultrafine particles, there are significant limits on how large an increase in ultrafine particle number would occur and how long such increase in ultrafine particles would last. More importantly, strategies developed for reducing accumulation mode mass will almost certainly involve a reduction in the rate of formation of ultrafine particles. EPA anticipates that the reduction in formation of ultrafine particles, as part of an overall PM control program, will cause a reduction in ultrafine particle number even though the accumulation mode mass is reduced. These points are developed more fully below.

As discussed in detail in the PM Criteria Document (U.S. EPA, 1996a; Chapters 3 and 6), ultrafine particles are condensed phase species with very low equilibrium vapor pressure, formed by nucleation of gas phase species. In the atmosphere there are three

---

[4]Accumulation mode particles generally extend from about 0.1 to as large as 1 to 3 μm in aerodynamic diameter. Ultrafine or nuclei-mode particles coagulate or grow by condensation into the accumulation mode. The fine mode consists of both (U.S. EPA, 1996b; p. IV-2-IV-3).

major classes of sources which yield particulate matter with equilibrium vapor pressures low enough to form ultrafine particles:

(1) Particles containing heavy metals. Ultrafine particles of metal oxides or other metal compounds are generated when metallic impurities in coal or oil are vaporized during combustion and the vapor undergoes nucleation. Metallic ultrafine particles are also formed from metals in oil or fuel additives that are vaporized during combustion of gasoline or diesel fuels.
(2) Elemental carbon or soot, C_e. C_e particles are formed primarily by condensation of C_2 molecules generated during the combustion process. Because C_e has a very low equilibrium vapor pressure, ultrafine C_e particles can nucleate even at high temperatures.
(3) Sulfates. Sulfuric acid (H_2SO_4), or its neutralization products with ammonia (NH_3), ammonium sulfate ($(NH_4)_2SO_4$) or ammonium acid sulfate (NH_4HSO_4), are generated in the atmosphere by conversion of sulfur dioxide (SO_2) to H_2SO_4. As H_2SO_4 is formed, it can either nucleate to form new ultrafine particles or it can condense on existing ultrafine or accumulation mode particles.

The concentration of ultrafine particles would be expected to increase with a decrease in accumulation mode mass, but to decrease with a decrease in the rate of generation of H_2SO_4. The rate of generation of H_2SO_4 depends on the concentration of SO_2 and OH, which is generated primarily by the photolysis of O_3. Thus, the reductions in SO_2 and O_3 that will form a major basis for attaining $PM_{2.5}$ and O_3 standards and implementation of Title II and Title IV Clean Air Act programs should lead to a decrease in the rate of generation of H_2SO_4 and a decrease in the concentration of ultrafine particles.

The commenter advances a theoretical argument but does not provide any quantitative assessment. In order to provide additional insight into the potential magnitude of the possible changes in particle numbers, EPA used a readily available aerosol formation model (Binkowski et al., 1996; Binkowski and Shankar, 1995) to illustrate the effects of changing accumulation mode mass and ultrafine particle generation rates. This aerosol dynamics model simulates the processes of nucleation, condensation, and coagulation of particles. The rate of condensation depends on the surface area available for condensation. The rate of coagulation depends on the number of particles and their sizes. Two representative situations have been examined, as described below. Details of the simulations are given in the accompanying Table.

The first simulation is similar to that posited by the commenter and addresses the first two classes of ultrafine particles outlined above. In essence, a pulse of ultrafine particles is added to an atmosphere with several different concentrations of accumulation mode particles. Only coagulation is modeled. This simulates the injection of auto exhaust into an air parcel moving across a busy street or highway or, with less precision, a plume of ultrafine particles near a stationary combustion source moving into an air parcel. The results of the simulations (cases 1-4 in the table) are shown in Figure 1. Simulations were performed for the injection of pulses of 5 and 2.5 $\mu g/m^3$ of ultrafine particles into an

air parcel containing 50, 25, or 12.5 μg/m^3 of accumulation mode particles. All simulations show the rapid coagulation of ultrafine particles into the accumulation mode, such that the majority of the pulse is gone in 20 to 40 minutes. The characteristic short lifespan of ultrafine particles, as illustrated in Figure 1, is one of the reasons they do not appear likely to be responsible for PM-mortality effects associations that have been observed in epidemiological studies of sensitive populations that spend the majority of the time indoors. (U.S. EPA, 1996; p. V-72 toV-73).

It is clear from the model results in Figure 1 that relatively large reductions in fine mass (e.g., from 50 to 25 μg/m^3; or from 25 to 12.5 μg/m^3) lead to only a small slowing of the disappearance of ultrafine pulses, and the increase at any time appears to be 10% or smaller. Although more significant effects might occur at much higher fine particle concentrations, given measured and estimated $PM_{2.5}$ air quality expected in U.S. cities, it is unlikely that implementing annual standards in the range of 15 μg/m^3 or 24-hour standards in the range of 65 μg/m^3 would produce fine mass changes as large as those simulated. Thus, these simulations indicate that, even in the unlikely event that fine particle strategies resulted in no reductions of directly emitted ultrafine particles, any effect on resultant ultrafine exposures would be small. To the extent that reductions in such ultrafine emissions do occur, the simulation results in Figure 1 (as illustrated in the reduced ultrafine pulse cases) show that the benefits of reducing the formation of ultrafine particles appear to be much greater than any effect of reduced accumulation mode mass.

In the second simulation, the formation rate of H_2SO_4 and the accumulation mode mass are varied and the three processes of nucleation, condensation, and coagulation are modeled (class 3 ultrafine particles as described above). This dynamic situation is more complex, and strongly dependent on relative sources and concentrations of fine particles, SO_2, ozone, relative humidity, and other factors, including conditions that favor more rapid transformation of SO_2 to sulfates. In this context, it is also important to note that one of the major strategies for reducing fine particle mass includes reduction of SO_2 emissions, which itself would tend to reduce the formation of ultrafine sulfate particles.

DESCRIPTION OF ULTRAFINE PARTICLE NUMBER MODEL AND CASES

A. INITIAL VALUES FOR PULSE SIMULATIONS FOR FIGURE 1

Background	**Case 1**	**Case 2**	**Case 3**	**Case 4**
Mass ultrafine, $\mu g/m^3$	0.25	0.25	0.25	0.25
Mass accumulation, $\mu g/m^3$	50	25	25	12.5
Number ultrafine, $\#/m^3$	3.91×10^{10}	3.91×10^{10}	3.91×10^{10}	3.91×10^{10}
Number accumulation, $\#/m^3$	1.78×10^{10}	8.90×10^{9}	8.90×10^{9}	4.45×10^{9}
Ultrafine Pulse				
Mass, $\mu g/m^3$	5	5	2.5	2.5
Number, $\#/m^3$	7.82×10^{11}	7.82×10^{11}	3.91×10^{11}	3.91×10^{11}

Note: All calculations done at one atmosphere, 295°K and 50% Relative Humidity. The geometric standard deviations for the ultrafine and accumulations modes are held constant at 1.7 and 2.0, respectively. The initial geometric mean diameters for the ultrafine and accumulation modes are 15 and 105 nanometers, respectively; however, these are allowed to vary with time. The background is assumed to contain ultrafine and accumulation mode, and the added pulse is assumed to be ultrafine mode. The aerosol dynamics model used is described in Binkowski and Shankar (1995) as modified by Binkowski et al. (1996).

B. INITIAL VALUES IN DYNAMIC SIMULATIONS FOR FIGURE 2

Initial Values:	**Case 6** (5@2)	**Case 7** (5@2)	**Case 8** (2.5@2)
Production Rate of Sulfuric Acid ($\mu g/m^3 s$)	1.11×10^{-4}	1.11×10^{-4}	5.57×10^{-5}
SO_4 ($\mu g/m^3$), ultrafine	0.25	0.25	0.25
SO_4 ($\mu g/m^3$), accumulation	25.00	12.50	12.50
Number ($\#/m^3$), ultrafine	3.91×10^{10}	3.91×10^{10}	3.91×10^{10}
Number ($\#/m^3$), accumulation	8.90×10^{9}	4.45×10^{9}	8.90×10^{9}

Note: All environmental conditions as well as the initial diameters and geometric standard deviations of modes as in A. above. The notation 5@2 denotes the SO_2 concentration in (5 ppb) and the conversion rate to sulfuric acid (2% per hour). Model described in: Binkowski and Shankar (1995) and Binkowski et al. (1996), as above.

INSERT FIGURES 1 AND 2 PAGE HERE

Figure 2 shows the variation in number of ultrafine particles for realistic variations in accumulation mode mass and H_2SO_4 production rates. With an accumulation mode mass of 25μg/m^3, an H_2SO_4 production rate of 2% h^{-1}, and 5 ppb of SO_2, the model predicts a slow increase in ultrafine particles with the concentrations raising from an initial value of 4 x 10^{10} particles/m^3 to 5.5 x 10^{10} after two hours (Case 6). If the accumulation mode mass is reduced to 12.5 μg/m^3, holding SO_2 concentration and oxidation rate constant, the number of ultrafine particles increases, reaching a maximum of 6.7 x 10^{10} at 45 min but then dropping to 4 x 10^{10} at two hours (Case 7). This suggests a factor of two reduction in non-sulfur related accumulation mode mass, unaccompanied by SO_2 reductions, could lead to a moderate increase in ultrafine particle number followed by a decline. However, if the fine mass reduction is at least partially attained by reducing SOx precursor emissions (as simulated in case 8 by reducing SO_2 from 5 to 2.5 ppb), or reductions in ozone and related precursors also occur, the formation rate of ultrafine sulfates slows and the total number is more likely to decrease with time due to coagulation of the ultrafine particles present in the assumed initial distribution. These simulations used a relative humidity of 50%. At higher humidities, characteristic of summertime photochemical episodes, the relative rate of ultrafine generation would be lower.

The following general conclusions can be drawn from an examination of the model output. For the same ultrafine particle or precursor input, a reduction in accumulation mode mass may lead to an increase in ultrafine particle number. However, at common concentrations in U.S. cities, the increase in ultrafine particles is much less than proportional to the decrease in accumulation mode mass. More importantly, reductions in ultrafine particle or precursor input dramatically reduce the number of ultrafine particles even when the accumulation mode mass is proportionally reduced.

Based on a consideration of the above factors, EPA concludes that a reduction in accumulation mode mass is more likely to be accompanied by a reduction in the generation rate of ultrafine particles so that the number of ultrafine particles will not increase. Even if there were to be a small increase in ultrafine particle number this would be offset by the public health benefits of a reduction in accumulation mode mass. The reasons EPA believes that mass, which is the metric correlated with effects in the epidemiological studies, is a more appropriate indicator than particle number, are discussed above. As noted in the Staff Paper, EPA believes that preliminary studies of the effects of ultrafine particles, including those of the commenter, suggest the potential for enhanced toxicity of this size range, and that further research in this area is of some importance. It is possible that freshly generated ultrafine particles relatively near significant sources could present an additional risk to health, above those associated with particle mass. It is also important to monitor particle number as well as mass to further delineate the relative effectiveness of strategies for reducing both particle mass and particle number.

iv. Consideration of separate/additional sulfate standard

Comment: Some commenters expressed the view that EPA should add a separate sulfate standard instead of $PM_{2.5}$ standards. Additional commenters (e.g., Resources for the Future, IV-D-2670; Colorado ALA and PIRG, IV-D-2095) expressed the view that EPA should add an additional sulfate standard to augment the $PM_{2.5}$ standards, while others (Ohio Edison, IV-D-2275) commented that EPA should exclude sulfates.

Response: See preamble to the final rule, sections II-B and II-C-1.

In the Staff Paper EPA noted that the most substantial laboratory and epidemiological data for any single class of fine particles exists for sulfates and associated acids (U.S. EPA 1996b, VII-14). The data for acids, which are more difficult to measure, is less consistent than for sulfates. Relatively strong correlations exist between acids, sulfates, and fine particles, making it difficult to single out any factor with confidence (U.S. EPA 1996a; p.13-93). Indeed, EPA considers sulfates useful as an indicator of fine particles for assessing the health effects literature. This literature suggests that reductions of regional sulfates as part of a fine particle control program would likely reduce mortality and morbidity risks for sensitive populations who reside in the East. For these reasons EPA concludes that it is not appropriate to exclude sulfates and that it is not appropriate at this time to establish a separate sulfate standard, alone or in combination with fine particle standards. Furthermore, EPA concludes that a sulfate standard, even if understood as an indicator of all fine particles, would be less likely to lead to controls of the other potentially harmful components of fine particles.

This information was presented to CASAC during EPA's review of the PM standards, and CASAC generally agreed with staff recommendations on this issue. The Committee advised the Administrator that "there is a consensus that retaining an annual PM_{10} NAAQS . . . is reasonable at this time" and that there is "also a consensus that a new $PM_{2.5}$ NAAQS be established" (Wolff, 1996b). Accordingly, EPA continues to conclude that it is appropriate to control fine particles as a group, as opposed to singling out particular components or classes of fine particles. The more qualitative scientific literature, evaluated in Chapter 11 of the Criteria Document and summarized in Section V.C of the Staff Paper, has reported various health effects associated with high concentrations of a number of fine particle components (e.g., sulfates, nitrates, organics, transition metals), alone or in some cases in combination with gases. Community epidemiological studies have found significant associations between fine particles or PM_{10} and health effects in various areas across the U.S. where such fine particle components correlate significantly with particulate mass. As noted above, it is not possible to rule out any one of these components as contributing to fine particle effects. Thus, EPA believes that the present data more readily support a standard based on the total mass of fine particles.

b. Averaging times

EPA received comparatively few public comments on the proposed averaging times. Those supporting $PM_{2.5}$ standards also strongly supported adopting both annual and 24-hour averaging times. Many of those opposing $PM_{2.5}$ standards, for the reasons discussed in Section II.B in the preamble to the final rule, provided contingent comments that variously supported both averaging times for $PM_{2.5}$ standards in the event the Administrator disagreed with their overall recommendations. Other opponents of $PM_{2.5}$ standards disagreed with having two standards on administrative grounds, or because some CASAC members did not support both averaging times.

Responses to comments on the relationship between standards for the two averaging times are presented in Section II.D of the preamble to the final rule and below. In essence, based on its examination of the effects data and air quality relationships, EPA believes that a single $PM_{2.5}$ standard (24-hour or annual) either would not provide adequate protection against effects of concern for all averaging times, or would be inefficient in the sense that it was more stringent than necessary for at least one averaging time. Contrary to commenters who focused on minority CASAC opinions, EPA notes that a clear majority of CASAC supported both 24-hour and annual standards. Of the 19 panel members who joined in the consensus for $PM_{2.5}$ standards, 17 (90%) recommended a 24-hour standard and 13 (70%) recommended an annual standard (Wolff, 1996b).

i. Annual standard as generally controlling standard

(1) *Comment:* Several Western state and local governments commented that the approach of using the annual standard as the controlling standard would not provide adequate protection in the Western U.S. where $PM_{2.5}$ emissions are seasonal in nature and characterized by short-term excursions. Other commenters (e.g., State of Washington, IV-D-7822) expressed the view that EPA's conclusion that a restrictive annual standard will control for high 24-hour levels is questionable because it is based on analysis of data from large, urban areas primarily in the east where emissions are relatively steady-state. These commenters asserted that evaluation of actual $PM_{2.5}$ data monitored in Spokane supports the claim that an annual standard would not effectively control for 24-hour events in the typical western setting.

Response: In its proposal, EPA specified the suite of PM standards such that the annual standard would be generally controlling, acknowledging that this would not be the case in every situation across the country because of the observed diversity of air quality distributions. Commenters are correct in observing that the annual standard is more likely to be controlling in areas with higher regional or wide urban area concentrations, such as is found in the Eastern U.S. and in Los Angeles. However, EPA maintains that the most appropriate risk management strategy, given the nature of the available epidemiological data, is to reduce area-wide population exposure and risk through a

generally controlling annual standard and protect against higher short-term peaks by an appropriate 24-hour standard.

As discussed in the preamble to the final rule, strategies for meeting a short-term standard focus on a characteristic "design value" episode responsible for peak concentrations. For PM, such peak values can be associated with single source contributions. Meteorology, relative source contributions, and resulting particle composition for that day may or may not be typical for the area or for the year. Yet the short-term exposure epidemiological results are largely drawn from studies that associated variations in area-wide effects with monitor(s) that gauged the variation in daily levels over the course of up to 9 years. The strength of the associations in these data is demonstrably in the numerous "typical" days in the upper to middle portion of the annual distribution, not on the peak days.[5] For these reasons, strategies that focus only on reducing peak days are less likely to achieve reduction of the mix and sources of urban and regional-scale PM pollution most strongly associated with health effects. Although designing control strategies to reduce annual levels may be more difficult than for 24-hour standards, the available short- and long-term exposure epidemiological data suggest it is also likely to result in a greater reduction in area-wide population exposure and risk.

(2) *Comment:* Some commenters questioned EPA's rationale for using an annual standard to protect against 24-hour effects because there is stronger evidence for a 24-hour standard from the more numerous short-term exposure studies. Others disagreed with the proposition that EPA's proposed approach would necessarily provide the most effective and efficient standards. In the view of some who opposed $PM_{2.5}$ standards, the likelihood that there are thresholds below which no effects occur means that a 24-hour standard would be more efficient than an annual standard. In this view, the reductions made on days that were below the threshold would provide no protection.

Response: EPA has fully responded to these comments in Section II.D. of the preamble to the final rule.

ii. Comments supporting alternative averaging times

(1) *Comment*: Some commenters (e.g., C.A.S.E., IV-D-4399) recommended shorter averaging times (e.g., 1-hour or 8-hour standards) to protect against short-term peak exposures.

---

[5] This point is buttressed by studies that have taken out a limited number of higher PM concentration days with little effect on the effects estimates or significance of the association (e.g., Schwartz et al., 1996; Pope and Dockery, 1992). One commenter (Bay Area Air Quality Management District IV-D-6502) provided an extended analysis of this kind for the Santa Clara County, California mortality-pollution data that finds the same result in this Western data set.

Response: See preamble Section II.D.1 for discussion of response to this comment.

(2) *Comment:* One commenter noted that an argument could be made for a multi-year standard based on the prospective cohort mortality studies. Another commenter recommended quarterly standards be set.

Response: As discussed in the preamble to the final rule, community epidemiological studies have reported associations of annual and multi-year average concentrations. The EPA has considered this evidence, which in short suggests that some health endpoints may reflect the cumulative effects of PM exposures over a number of years. In such cases, an annual standard would provide effective protection against persistent long-term (several years) exposures to PM. EPA notes that the form of the annual standard requires three years of data. Requiring a much longer averaging time would also complicate and unnecessarily delay control strategies and attainment decisions.

EPA also explicitly considered the suggestion of using a quarterly (i.e., 3-month) average, in view of the observed seasonality of concentrations of fine particles and their precursors in some areas (e.g., wintertime smoke from residential wood combustion, summertime regional acid sulfate and ozone formation). However, different seasons are likely of concern in different parts of the country, and the current evidence does not provide as satisfactory a quantitative basis for setting a national fine particle standard in terms of a seasonal averaging time as for an annual standard, in combination with a 24-hour standard. Ultimately, EPA rejected a seasonal averaging time in this review, focusing the 24-hour standards on protecting against seasonal excursions not adequately addressed by the annual standard.

c. Standard levels

i. Characterization of CASAC views regarding $PM_{2.5}$ standard levels

Comment: Some commenters expressed the view that because CASAC did not agree on specific levels and averaging times for the standards, it would be inappropriate to establish any standard. Some also provided various characterizations of CASAC opinions, for example by combining the views of panel members supporting standard levels higher than those EPA proposed with the views of panelists who chose not to recommend levels as constituting a majority that did not support EPA's proposals. These commenters also noted that only one CASAC panel member recommended annual standards as low as EPA proposed. Additional comments stated that CASAC did not endorse EPA staff's recommended ranges in the Staff Paper.

Response: Before responding to the specific points raised by these commenters, EPA notes that some of them appear to rest on questionable assumptions about the role and purposes of CASAC review. Briefly stated, Congress expected CASAC to advise the Administrator on the scientific basis for NAAQS decisions and to recommend such

revisions in the air quality criteria and NAAQS as it considers appropriate. The Administrator, in turn, must consider CASAC's advice and recommendations but is not bound by them. There is no requirement that there be a consensus on scientific issues before the Administrator may revise standards or establish new ones. See, e.g., Lead Industries Ass'n v. EPA, 647 F.2d 1130, 1154-55 and n.51, 1160-61 (D.C. Cir. 1980). Indeed, as discussed in section IV.A of the preamble to the final rule, uncertainty and controversy on scientific issues are inherent in the statutory scheme, which in effect requires decisions "at the very 'frontiers of scientific knowledge'" where "disagreement among the experts is inevitable." 647 F.2d at 1147, 1160. Section 109(d)(1) of the Act nonetheless requires the Administrator to review the criteria and NAAQS at least once every five years, to decide whether revisions are appropriate, and, if so, to make the appropriate revisions.

By statute as well as historical practice, CASAC includes representatives of a variety of disciplines. As a result, individual panel members differ in the expertise they bring to particular scientific issues, and individual members often choose not to express opinions on matters outside their own areas of expertise. CASAC has also been careful to distinguish between the advice and recommendations it provides as a Committee and the views expressed, orally or in writing, by individual panel members. In addition, CASAC typically acknowledges, as it did in this review (Wolff 1996), that NAAQS decisions require the Administrator to make public health policy judgments as well as determinations of a strictly scientific nature. Since the Committee began advising EPA in the late 1970's, in fact, it has generally stopped short of recommending specific standard levels, as opposed to advising on ranges of possible levels, developed by EPA staff, for the Administrator's consideration.

Thus, the lack of a consensus among CASAC panelists on such matters as specific levels and averaging times is clearly no bar to revising existing standards. Indeed, the lack of a consensus would not excuse the Administrator from her statutory duty to determine whether revisions are appropriate and, if so, to make the appropriate revisions. In this context, counting the "votes" of individual CASAC panelists on such issues as standard levels is of interest and can be illuminating, but it is ultimately less important than careful consideration of the substance of the Committee's advice on the underlying scientific issues.

In any case, the characterizations of CASAC recommendations and guidance given by many of these commenters do not reflect an appropriate and complete summary of the advice given by the panel. Indeed, given the variety of opinions expressed by individual panelists and the variety of summaries submitted by various sources, it is clear that any simple summary of the advice given by individual panelists is subject to question. For these reasons, EPA believes it is important to examine the advice and recommendations of the committee *per se* before attempting to consider the views of individual panelists.

In its March and June letters, CASAC provided general closure statements indicating that both the Criteria Document and the Staff Paper, which contains conclusions and recommendations on standards, were scientifically adequate for use in the regulatory decision-making process. Individual members made comments regarding the recommended staff ranges and, in response to the Chairman's request, many gave their own opinions either on ranges or on specific standards they would recommend. However, the Committee made no attempt to discuss or put forward any specific consensus recommendation on a range or specific standard levels, nor to comment as a group on the staff range. The summary table in the Committee's June 13, 1996, letter reflects the diversity of the specific opinions registered by the individual members.

As indicated above, the lack of convergence with respect to opinions on specific standards by the 11 individual members who offered them does not relieve the Administrator of the responsibility to make a decision on the need to revise the standards, about which there was a clear CASAC consensus, or the need to select specific averaging times, forms, and levels for revised standards. The 8 panelists who chose not to recommend specific ranges cannot be counted as not supporting adoption of specific $PM_{2.5}$ standards at this time. Instead, they left this judgment to the Administrator, guided by their comments on the staff ranges. While it appears that *any* standard the Administrator could choose will be at variance with the opinion of one or more CASAC members, a decision is nevertheless required. Moreover, the spread of opinion in this case (e.g., support for 24-hour standards ranging from 20 to 75 μg/m3) is actually smaller than the range CASAC recommended for the 24-hour PM_{10} standard in 1982 (150 - 350 μg/m3) (Friedlander, 1982). The fact that none of the current CASAC PM panel recommended relaxing the level of protection afforded by the current PM standards indicates a clear consensus for more stringent standards than was evidenced in the earlier review.

A fair reading of the opinions expressed on the averaging time for $PM_{2.5}$ standards is that, of the 19 panel members who joined in the consensus for $PM_{2.5}$ standards, 17 (90%) recommended a 24-hour standard and 13 (70%) recommended an annual standard. This clearly supports the Administrator's proposal to adopt both averaging times. All members speaking to the issue of form recommended a more robust form, such as that adopted in the final decision. The largest divergence of opinion was on the levels for the standards, and on the relative protection afforded by 24-hour as opposed to annual standards.

EPA believes that most of the CASAC panelists' opinions on averaging times reflected their judgments on the greater relative strength of the short-term exposure epidemiological studies, a judgment that EPA shares. Based on the recommendations of staff and some CASAC members, the Administrator proposed to use the annual standard to provide protection against both short- and long-term exposures. This led to a tighter level for the proposed annual standard relative to the proposed 24-hour standard. EPA believes this approach is neither inconsistent with the underlying science nor discordant

with the overall advice of CASAC. In evaluating these recommendations, it is also important to note that panelists were reacting to staff recommendations for levels that were expressed in a single exceedance form (24-hour) with no spatial averaging; for similar levels, these forms are clearly more stringent than the forms proposed and adopted for the final standards. Indeed, when the projected relative protection afforded by the combination of $PM_{2.5}$ standards originally proposed by EPA is compared with the combined 24-hour and annual opinions of each of the 11 CASAC panelists, the protection afforded by the proposed and final standards (in terms of the estimated number of people living in areas that need improved air quality) is toward the middle portion of that estimated for the combined 24-hour and annual range of recommendations by the 11 individual CASAC panelists. Among those recommending ranges of standards including combined levels generally tighter than or close to those proposed by EPA were two epidemiologists, a health expert, a toxicologist, and an air quality expert. By contrast, only two of the six recommending less protective standards were health scientists. In short, a simple comparison of CASAC panelists' June 1996 opinions on standard levels with EPA's proposal or final rules cannot convey the full extent to which the Administrator incorporated the panel's scientific judgments as well as the personal views of individual panel members in reaching her final decision on the proposal.

With respect to specific summaries of CASAC opinions on standard levels presented by some commenters, EPA does not think it is appropriate to combine the recommendations of members who expressed personal opinions on the standard levels with those who stated that they preferred not to express such opinions. For example, with respect to the 24-hour standard, the table of opinions in the CASAC closure letter shows four panelists who recommended specific 24-hour standards at or above the upper bound of the range recommended by EPA staff, and four who recommended standard ranges that extended well below that adopted by the Administrator. As noted above, EPA considered the results of the daily studies in reaching a decision on the annual standard. It is therefore appropriate to note that the final annual standard achieves a level of protection that falls in the upper portion of the combined range of recommended 24-hour and annual levels put forth by this latter group of panelists. EPA recognizes that some panel members who did not express an opinion on any specific standard levels did comment with respect to the upper end and lower portions of the staff range. For example, two such panelists commented that the upper portion of the 24-hour standard may be too low. In recommending that the upper portion of the range be widened, these two particular panelists did not recommend that the staff modify the lower bound of the range, and it would be appropriate to conclude that they believed the science supported consideration of 24-hour standard levels in a relatively wide range that extended well below that proposed or adopted in the final decision.

As indicated above, it is important to separate the personal opinions that individual members might express on particular policy choices such as standard levels from their scientific conclusions on the range of options that is supported by the science and should be considered by the Administrator. It is also important to recognize that individual

members who chose not to offer their own specific options or ranges were not necessarily in disagreement with the ranges put forward by EPA staff or opposed to selection of a particular set of numbers by the Administrator. Indeed, it would be illogical to recommend the adoption of $PM_{2.5}$ standards without recognizing the responsibility of the Administrator to decide upon levels that are consistent with the underlying science in the criteria and the legal mandates to protect public health that are inherent in the requirements of the Clean Air Act.

Speaking as a committee, CASAC clearly recommended that the current PM standards be revised by establishing $PM_{2.5}$ standards at the conclusion of the present review, and that -- as stated prominently in the Staff Paper -- the data are adequate to support levels for both 24-hour and annual standards. The lack of a consensus on particular levels in no way invalidates the overall conclusions that new standards should be established nor the need for the Administrator to come to specific conclusions in discharging her statutory duty.

ii. Consideration of "equivalent" levels and other levels above those proposed

Comment: Numerous commenters opposed revised standards but provided "contingent" comments that recommended levels well above those proposed by the Administrator. These commenters placed great weight on (1) factors outlined elsewhere in the Summary of Comments that led them to oppose any revisions to the PM standards, including the uncertainties and limitations in the available health effects studies considered individually, such as the possible existence of effects thresholds and unanswered questions regarding the causal agent(s) responsible for the reported health effects; and (2) the limited amount of research currently available that has measured $PM_{2.5}$ directly. A substantial group recommended that $PM_{2.5}$ standards be selected so as to be equivalent to or close in stringency to the current PM_{10} standards, and cited the opinions of some CASAC PM panel members as support. Some of these commenters provided supplemental analyses of air quality data, arguing that they demonstrate that "equivalent" standards would be at $PM_{2.5}$ levels higher than the highest proposed (65 $\mu g/m^3$, 24-hour average and 20 $\mu g/m^3$ annual average), reaching as high as approximately 95 $\mu g/m^3$ 24-hour average and 27 $\mu g/m^3$ annual average.

Response: Having evaluated these comments, the EPA rejects both their underlying rationale and the specific recommendations for $PM_{2.5}$ standard levels that result in similar or only marginally more protection than that afforded by the current PM_{10} standards. In the first instance, there is no scientific basis for complete "equivalence" when one measure (PM_{10}) contains coarse particles and the other ($PM_{2.5}$) does not. As CASAC recognized in its review, the wide variability in $PM_{2.5}/PM_{10}$ ratios in time and location precludes defining uniform $PM_{2.5}$ standards that would provide close to "equivalent" protection to the current standard in all or even most areas. Aside from such technical problems, which are inherent in the commenters' supporting analyses on the issue of defining "equivalent" standards, this approach would be inconsistent with the Administrator's conclusions regarding the adequacy of the current standards and the need

to provide additional protection as articulated in the preamble to the final rule. The EPA believes that, despite well recognized uncertainties, the consistency and coherence of the epidemiological evidence and the seriousness of the health effects require a more protective response than provided by "equivalence" or a marginal strengthening of the standards. Moreover, EPA believes that the standard levels should be based on the most recent assessment of the scientific criteria for PM, not on applying uncertain ratios to standard decisions based on much more limited evidence in 1987.

The EPA also rejects the premise of some who suggest that adopting standards that prompt little or no additional control would cause no delay in risk reduction as compared to conducting monitoring and research now and setting more stringent standards after the next review. These comments do not consider the realities of implementing air quality standards, which ensure that such an approach would add several years to the risk reduction process. Thus, aside from her obligations under the statute, the EPA believes that the most prudent and appropriate course is to establish appropriately protective standards now that put into motion monitoring, and strategy development programs, while at the same time pursuing an expanded research program to improve implementation and to inform the next periodic review of the criteria and standards.

iii. Proposed levels inappropriate

(1) *Comment:* Several commenters (e.g., Transportation Corridor Agencies, IV-D-2533; Ford Motor Co., IV-D-5323) objected to the specific levels EPA proposed for the standards on the basis that the proposed annual and/or 24-hour levels approach or are below "background" levels in non-urban areas. Some objected to EPA's use of pristine and high elevation sites to estimate background levels.

Response: In any discussion of "background" PM, it is important to define what is included. With the air pollution community, the term "background" is sometimes used in reference to uncontrollable pollution levels from "natural" sources, and sometimes used to mean the observed concentrations in unpopulated or rural areas. In the latter case, the term includes controllable pollutant levels derived from regional anthropogenic as well as natural sources. For clarity, CASAC advised that the Criteria Document and Staff Paper define "background PM" as "the distribution of PM concentrations that would be observed in the U.S. in the absence of anthropogenic emissions of PM and precursor emissions of VOC, NO_x, and SO_x in North America" (U.S. EPA 1996b; p IV-12). In essence, this fraction represents that portion of PM that cannot be reduced by controlling U.S. or other anthropogenic sources on the continent.

EPA disagrees that its annual and/or 24-hour standards approach or are below the estimated background concentrations for $PM_{2.5}$ that are based on the assessment contained in the Criteria Document and Staff Paper, and reviewed by CASAC. The estimated annual average regional background $PM_{2.5}$ levels are 2 to 5 $\mu g/m^3$ in the East, and 1 to 4 $\mu g/m^3$ in the West (Staff Paper, Table IV-3). The highest background

estimates values are therefore about 3 to 4 times lower than the proposed annual $PM_{2.5}$ standard. Absent forest fires, which are exempted under EPA's Natural Events Policy, peak background 24-hour $PM_{2.5}$ values nationwide are expected to be lower than 15 to 20 $\mu g/m^3$. When accounting for the exceedances allowed under the 98th percentile form and level of the 24-hour standard, these background levels are about 4 times or more lower than allowed by either the proposed or final standards.

This assessment of background levels as defined in the Staff Paper was not predicated solely on the observed levels at high elevation sites as suggested by commenters, but on published evaluations of the composition of observed levels in non-urban and other background locations around the country. The fact that annual levels as high as 7 to 13 µg/m3 are observed in some non-urban areas is generally reflective of either high regional anthropogenic contributions (e.g., in the Eastern U.S.) or the influence of local or urban scale sources on rural areas (e.g., the California Desert and some other western locations). Unlike natural or extra-continental background contributions, these sources are controllable.

In summary, although estimates of natural emissions and concentrations are uncertain, the available scientific evidence indicates that the proposed annual and 24-hour $PM_{2.5}$ standards are well above those that would occur under typical background conditions. The only exceptions to this involve short-term exceptional or episodic events (e.g,. forest fires, volcanic eruptions) that are routinely exempted in implementation policy under the current PM_{10} standards.

EPA agrees that regionally transported anthropogenic PM levels may approach or form a significant portion of the standard levels. Indeed, the annual standard is expected to result in region-wide reductions in fine particle concentrations. The fact that standards may require additional reductions from controllable pollution sources is not, however, an appropriate basis for selecting a level higher than that requisite to protect public health. Such regional or urban-scale reductions should lead to significant population risk reductions, because they affect large numbers of individuals across wide areas.

(2) *Comment:* One commenter (TVA, IV-D-2289), who generally opposed establishing any fine particle standard, recommended setting a $PM_{2.5}$ standard at a level equivalent to existing $PM_{2.5}$ concentrations, albeit varying across the country. Another commenter (Resources for the Future, IV-D-2670) recommended consideration of a $PM_{2.5}$ standard for the East and a $PM_{1.0}$ standard for the West of similar stringency. These commenters note that legal authority for their suggestions does not currently exist.

Response: See earlier response with respect to the concept of "equivalence" as a basis for standard setting in this review. With respect to varying standards by region, EPA is required by law to establish a uniform national standard to protect public health; thus, such options are not appropriate for consideration. To the extent that regional differences

in PM air quality exist, it is appropriate to consider accommodating them in implementing the standards.

(3) *Comment:* Some commenters who advanced arguments in favor of "equivalent" standards (see previous subsection) presented or relied on technical analyses of $PM_{2.5}/PM_{10}$ ratios to suggest that the upper bound of the ranges of standard levels presented in the staff paper were more stringent than EPA suggested, and further that this biased the selection of the proposed levels. AISI (IV-D-2242), in particular, also argued that EPA's approach to predicting $PM_{2.5}$ concentrations understated the estimated number of new non-attainment areas, and presented alternative estimates.

Response: As noted above, EPA does not believe the $PM_{2.5}$ standards should be based on "equivalence" with the current standards. EPA has addressed commenters' claims related to EPA's methodology to predict $PM_{2.5}$ values from the larger PM_{10} dataset in a staff memorandum (Fitz-Simons, 1997). As discussed therein and summarized below, EPA disagrees with commenters' claim that EPA data sets are biased towards rural sites and that EPA's predictive model is inherently biased. EPA staff did not ignore evidence submitted by commenters regarding the range of equivalence in various regions of the country in developing the staff range. The Staff Paper clearly states: "in some Eastern areas, a $PM_{2.5}$ level as high as about 100 $\mu g/m^3$ could correspond to the current 24-hour PM_{10} standard level, whereas in some Western areas the corresponding $PM_{2.5}$ level could be as low as about 50 $\mu g/m^3$. Thus, there is no 'equivalent' level that applies nationally based on information on ratios between $PM_{2.5}$ and PM_{10} " (Staff Paper, p. VII-24). Given that the level of the final 24-hour standard of 65 $\mu g/m^3$ is at the high end of the staff range, commenter's suggestion that the staff recommendation biased the decision to a still lower number is clearly without merit. Similarly, EPA clearly based the level of the annual standard on the available scientific evidence, without reference to equivalence with the current standards.

EPA also disagrees with the less germane comments regarding bias in EPA's projection methodology. EPA staff developed the methodology at the request of the CASAC, and the air quality experts on the panel, including the chairman, reviewed the underlying data bases, the approach and initial the results as presented in the Staff Paper. EPA has clearly noted the substantial uncertainties associated with such projections, but does not believe commenters have demonstrated inherent bias. The AISI comment that the underlying data base has a rural bias is without merit and inconsistent with the conclusions of the organization's expert consultant. Although the initial database developed by staff earlier in the standards review process may have over-represented rural sites, the version used in the Staff Paper and as the basis for EPA's projections and subsequent analyses is more balanced. The commenter's own consultant acknowledges this in an attachment to the comments submitted to EPA to support this claim (Attachment B of AISI, Cooper Environmental Services), stating that the EPA database "is probably seasonally balanced and is now more geographically balanced because California data have been added."

Similarly, EPA believes that its predictive model developed from the dataset is uncertain, but unbiased. A review of the detailed comments finds that the analysis provided by commenters uses an inappropriate approach for assessing model bias (Fitz-Simons, 1997). Commenters' second basis for claiming the model is biased, namely because it underpredicts the number of non-attainment counties relative to commenters's estimates is also fundamentally flawed. Commenter relied on an assessment that actually used the regional $PM_{2.5}/PM_{10}$ ratios that were in EPA's data base. The increased number of counties projected was not related to model bias, but largely to the decision to include all counties in an SMSA if any county showed a predicted violation. This is an implementation policy assumption, not a model bias. EPA believes that, while its model provides only moderately accurate predictions given the differences in sources of fine and coarse particles across the country, the model does not systematically overpredict or underpredict $PM_{2.5}$ values.

(3) *Comment:* Colorado PIRG and Colorado ALA (IV-D-2095) claim that the Agency did not consider the magnitude of risk at the proposed level of the annual standard of 15 $\mu g/m^3$. The commenter cites a per capita risk analysis attatched as support for a more stringent level of the annual standard.

Response: EPA relied upon the scientific information in the Criteria Document and Staff Paper (not the RIA as the commenter also claims) in selecting the level of the NAAQS. In these documents as well as in the preamble to the proposed and final rules, EPA acknowledges that some risk of effect is seen across a substantial range of PM concentrations, including levels approaching the lowest in the scientific studies. Moreover, EPA specifically recognizes that its NAAQS are not risk-free but are designed to reduce risk sufficiently to protect the public health with an adequate margin of safety (see Section III.A.2). EPA has considered both the strengths and the limitations of the available evidence as well as alternative interpretations of the scientific evidence, as discussed above and in the Appendices. As illustrated in Figure 2 of the proposal, the risk at concentrations approaching 15 $\mu g/m^3$ is uncertain, and depend greatly on assumptions about the shape of the concentration-response function (Figure 2c). These increasing uncertainties limit the confidence that can be placed in the risk estimates at lower concentrations.

EPA examined the per capita risk estimates provided by commenters, and finds they are based on a questionable premise. The calculation follows an approach often used to estimate expected incidence for life time exposure to carcinogens. While it is certainly true that the relative risk in a single day may greatly understate the risk of exposure to PM in sensitive populations, this approach of estimating the risk of mortality carries the methology too far. The available studies used to provide the estimate were among those that found the risk of mortality is primarily in the elderly (> 65 years) and those with pre-existing cardiopulmonary disease. As a result, the explicit calculation of a cumulative risk that assumes those individuals would live another 70 years (e.g., to age 135) is not

reasonable. Moreover, the upper bound here appears inconsistent with long-term studies of cumulative PM risk (Pope et al., 1995). While the example provided by commenters may provide an interesting illustration, EPA does not agree that it represents a reliable estimate of risk.

iv. Consideration of more stringent levels

Many commenters, notably environmental groups and some States, strongly supported standard levels more stringent than those proposed by EPA. These commenters supported EPA's conclusions regarding the epidemiological studies but would place much less weight on uncertainties related to the concentration-response relationships for $PM_{2.5}$ as a surrogate for PM and the relative importance of various PM components. Based on their evaluation of the information, and citing the support of some CASAC panel members, these commenters variously recommended 24-hour $PM_{2.5}$ standards as low as 18 to 20 $\mu g/m^3$ and annual standards of 10 to 12 $\mu g/m^3$.

As discussed in section II.F of the preamble to the final rule, such standards would result in commensurate reductions in health risks only if, in fact, there is a continuum of health risks down to the lower end of the ranges of air quality observed in the key epidemiological studies, and only if the reported associations are, in fact, causally related to $PM_{2.5}$ at the lowest concentrations measured. Setting standards at low levels where the possibility of effects thresholds is greater, and where there is greater potential that other elements in the air pollution mix (or some subset of particles within the fine fraction) become more responsible for (or modify) the effects being causally attributed to $PM_{2.5}$, might result in regulatory programs going beyond those that are needed to effectively reduce risks to public health. While placing substantial weight on the results of the key health studies in the higher range of concentrations observed, EPA is persuaded that the inherent scientific uncertainties are too great to support standards based on the lowest concentrations measured in such studies, which approach the maximum range of $PM_{2.5}$ values estimated for short-term background conditions.

EPA notes that the range of levels recommended by these commenters for a 24-hour $PM_{2.5}$ standard is close to the lower bound levels recommended by four CASAC panel members (20 $\mu g/m^3$); however, no CASAC panelists supported an annual $PM_{2.5}$ standard as low as 10 to 12 $\mu g/m^3$. For reasons specified in the preamble to the final rule, EPA chose to use the annual standard, rather than the 24-hour standard, to be generally "controlling" with respect to health protection provided by both standards. The final annual standard level is at the lowest level recommended by any CASAC member, and the 24-hour standard is at the upper end of the range recommended by the EPA staff and most CASAC members. Taken together, these standards provide protection that is substantially greater than afforded by the current standards, well within the ranges recommended by CASAC and EPA staff.

The approach taken by EPA focuses primarily on standard levels designed to limit annual $PM_{2.5}$ concentrations to levels somewhat below those where the body of epidemiological evidence is most consistent and coherent, in recognition of both the strengths and the limitations

of the full range of scientific and technical information on the health effects of PM, as well as associated uncertainties, as interpreted by the Criteria Document, Staff Paper, and CASAC. The Administrator believes that this approach appropriately reflects the weight of the evidence as a whole.

(3) *Comment:* Some commenters (Colorado PIRG IV-D-2095, NRDC IV-D-2267) argued that EPA should have followed the methodology it used to assess the epidemiological evidence for the purpose of establishing levels for the 1987 PM_{10} NAAQS. In that approach, EPA developed estimates for concentrations at which effects were "likely," "possible," and "no effects likely." Selection of EPA's level is based on the risk assessment and primarily on the long-term exposure studies with support from short-term exposure studies. Both Dockery et al. (1993) and Pope et al. (1995) indicate that mortality and morbidity effects continue down to exposures as low as 9 $\mu g/m^3$.

Response: The approach followed in the 1987 standard review is not fundamentally different than that used in the current review. The 1987 approach, which also involved converting concentrations from metrics such as British Smoke and TSP to PM_{10}, was drawn directly from the available scientific information as it was summarized in the Criteria Document and Addendum. Because that information was dominated by studies of high concentrations, particularly in the London episodes, the Criteria Document placed some emphasis on estimated concentrations at which PM effects were thought to be "likely," a term viewed by some as demonstrated effects levels at the concentrations involved in such studies. At the concentrations recorded in such studies it was possible to observe increased daily mortality on single episode days. There was little dispute among CASAC members and other scientists regarding causality at high concentrations, and the distinction between "likely" and "possible" was more relevant to levels then recently observed in the U.S. The remaining examination focused on the range where effects were possible, and lower levels where available evidence suggested no effects. The range of possible effects extended to the lowest levels observed in 14 London winters, which still reflected significant urban PM pollution. It was these levels that formed the basis for the lower end of the EPA staff and CASAC range of interest for standard setting. EPA concluded that, although the risk of mortality suggested by these London data extended to levels below those selected for the PM_{10} standards, the evidence was not sufficient to warrant more stringent levels.

In the present review, the Criteria Document focused on the numerous recent studies providing evidence that PM related effects occur at or below those allowed by the current standards. At these lower levels, it is not possible to readily detect unequivocal increases in mortality during single days, as in the early London data, and analyses in particular urban areas require months or years of daily data to reach conclusions regarding increased risk. The lowest observed levels for daily studies overlap those of background conditions, not polluted urban atmospheres. Therefore, the emphasis of this review has been on the likelihood of effects at levels that were predominantly below the level of the current standards down to the lowest observed values. As evidenced by the spread of

opinions among CASAC panelists, within that range, there is no general agreement on demonstrated effects levels; PM effects at levels at or below the current standards are recognized as at least possible by most, and judged likely by some.

The Criteria Document concludes that the observed PM effects-associations at levels at and somewhat below the current standards suggest a "likely" causal relationship. This cannot be interpreted to mean a demonstrated effects level. Moreover, unlike the London situation discussed above, it cannot be said either that effects are likely or even that the evidence indicates they are possible at the lowest observed levels, making consideration of such levels inappropriate for standard setting. For this reason, the lower bounds of staff recommended standard ranges were well above these levels. As discussed more fully in Section II.F of the preamble to the final rule, EPA believes the key uncertainties with respect to the available scientific evidence on PM at such levels are such that the standard should focus on the strongest data, which are at and above the mean levels observed in the short-term exposure studies. The likelihood of significant health risks, if any, becomes smaller as concentrations extent below these levels. In the end, the Administrator has selected standards based on consideration of the weight of the scientific evidence, recognizing that some risk of effects remains at still lower levels. In this respect, the approach followed in reaching this decision is completely consistent with that followed in the previous review.

(2) *Comment:* One commenter (Sonoma County APCD IV-D-7013) requested that EPA explain why short-term exposure health effects at 35 $\mu g/m^3$ are not significant enough to be prevented and why the proposed standard was not set to at least prevent health effects in the 40-50 $\mu g/m^3$ range.

Response: EPA has considered the weight of the scientific evidence including uncertainties in selecting the final standards, recognizing that some risk remains at still lower levels. This comment focuses only on the 24-hour standard level. By contrast, EPA selected an approach which considers the combined protection afforded by both the annual and 24-hour standards and treats the annual standard as generally controlling. As discussed above and in Section II.F of the preamble to the final regulation, EPA believes the key uncertainties with respect to the available scientific evidence on PM are such that the standard should focus on the strongest data. EPA considers the strongest evidence from the daily studies to be at concentrations toward the middle of the distribution of annual concentrations. Accordingly, the annual standard would serve to reduce short-term exposures in the range specifically raised by commenters in areas where such levels occur with enough frequency to cause annual levels above 15 $\mu g/m^3$, either in particular urban areas or in regions with elevated fine particle levels. EPA does not believe that the scientific evidence on the cumulative risk of very infrequent excursions at such levels is certain enough to warrant the additional protection of a more stringent 24-hour standard.

(3) *Comment:* Some commenters supported annual $PM_{2.5}$ standard levels below levels recommended in the Staff Paper. They cite Staff Paper discussions of lower limit of detection or interpret lowest observed exposures in long-term exposure studies (e.g., noting that in Dockery *et al*. (1993) and Pope *et al*. (1995), mortality and morbidity effects continue down to exposures as low as 9 $\mu g/m^3$).

Response: As stated in the Staff Paper, the purpose of discussing lower limits of detection was to support sensitivity analyses performed in developing ranges of estimated risks. One of the elements needed for these sensitivity analyses is alternative PM concentration ranges over which reported concentration-response functions would be applied. The lower limits of detection would serve as the lower end of such ranges (Staff Paper, Appendix E). Indeed, staff recommendations regarding ranges for consideration did not extend this low due to the increasing uncertainties at lower levels such as the possibility of a threshold. The lower end of the staff recommended range in the Staff Paper for the annual standard was not related to and well above the lower limit of detection. The Staff Paper also notes that the central tendency (i.e., mean or median) concentration is generally the point for which there is the greatest confidence. As discussed above and more fully in Section II.F of the preamble to the final rule, EPA believes the key uncertainties with respect to the available scientific evidence are such that the standard should focus on the strongest data, which are at and above the mean levels observed in the short-term exposure studies. For reasons discussed in Section II.F of the preamble to the final rule, the level of the annual standard is based primarily on the annual levels observed in the short-term epidemiological studies; in the Administrator's judgment, the long-term exposure studies, which are subject to greater uncertainties, are consistent with the level she selected.

EPA also notes that the lowest fine particulate concentration reported in Pope *et al*., (1995) is a median value of 9.0 $\mu g/m^3$, not a mean value. The lowest mean value would be expected to be about 11 $\mu g/m^3$, based on Freas (1997) analysis of typical ratios related to $PM_{2.5}$ mean/median ratios (1.21). This is close to the lowest mean $PM_{2.5}$ level observed in the 6-city study (Dockery *et al*., 1993). This raises the levels underlying commenters' rationale to about as low as 11 $\mu g/m^3$, but has no implications for EPA's rationale, which gives primacy to the 24-hour studies, and uses the annual results in a supportive role in evaluating the margin of safety. EPA believes commenters' approach does not appropriately recognize the fundamental uncertainties regarding the risks at such lower concentrations. In the end, EPA has selected standards that protect public health, based on consideration of the weight of the scientific evidence.

(4) *Comment:* In partial support of the argument that $PM_{2.5}$ standards should be more stringent than those proposed by EPA, the Bay Area Air Quality Management District (IV-D-6502) submitted a new analysis of PM and mortality in Santa Clara County. This analysis finds that "a large number of premature deaths would still occur in the San Francisco Bay Area, even under attainment of the new standards" (commenter's Attachment 1). In addition, a brief report from a new study is attached that finds "a

reduction in the proposed $PM_{2.5}$ 24-hour standard from 50 $\mu g/m^3$ to 30 $\mu g/m^3$ would result in a reduction of as many as 1300 deaths per year in the Bay Area" (commenter's Attachment 2).

Response: A provisional examination of these supplemental analyses of the Santa Clara data indicates that they appear to offer further evidence for health effects of $PM_{2.5}$ at low concentrations. In the first attachment the data set from a previous publication (Fairley, 1990) was divided into subsets that excluded days with higher PM levels. As discussed in the preamble, the HEI reanalysis of the earlier publication, using COH as the PM indicator, found essentially the same results as reported by the original investigator. The report in Attachment 1 adds two new facets to the original investigation, a conversion of COH to $PM_{2.5}$ and an examination of effects when data from high-exposure days are excluded for analysis. EPA has some concern about the methods used to extrapolate $PM_{2.5}$ values from COH measurements, and the validity of the resulting $PM_{2.5}$ estimates. As stated in the Criteria Document (p. 4-55), "Any attempt to relate COHs to $\mu g/m^3$ requires site-specific calibration of COH readings against mass measurements determined by a collocated gravimetric device, but the accuracy of such mass estimates are still subject to question." In this case, data from the same location (1980-1986) were used, but data collected at a later time (1990 and later) were used in the calibration of COH against $PM_{2.5}$, and PM levels had been reduced over this time period. This has clear implications for the applicability of the calibration.

It must also be recognized that the findings of these two analyses have not been subject to peer review or submitted for publication. As has been stated previously, the EPA's decision must be based on studies evaluated in the criteria review process, including review by CASAC. Were these analyses to be considered in this rulemaking process, it appears that they would offer evidence for health effects related to PM exposures in an area that is in attainment with the current PM standards. This is particularly true for the findings presented in the second attachment, which is a completely new study (though it is based on the same air pollutant data set) that expanded the analysis to the full year, and extends predicted mortality reductions to surrounding Bay Area counties. Although it appears unlikely that these analyses, standing alone, would materially change EPA's conclusions if considered in this rulemaking, EPA encourages the commenters to submit these studies for publication so that they may be considered in the next review of the PM standards.

d. Forms of the standards

A broad range of public comments were received in this area. The comments regarding the form of the annual $PM_{2.5}$ standard centered mainly on spatial averaging provisions. The comments on the form of the 24-hour standard focused on the adequacy of health protection, particularly for western areas, and/or comparisons with the 1-expected-exceedance form. The comments are summarized primarily in section II.A.2.a.(4) of the Summary of Comments and significant comments are highlighted in the preamble to the final rule. In addition to the

responses contained in the preamble to the final rule in sections II-B and II-E, EPA provides the following additional responses to specific issues.

i. Spatial averaging

(1) *Comment:* A number of commenters objected to the concept of population-oriented monitors and expressed the view that any monitor regardless of where it was sited should be eligible for comparison to the annual $PM_{2.5}$ standard. They further maintained that the proposed provisions for spatial averaging would fail to provide adequate health protection because "clean areas" and "dirty areas" would be averaged together. Some commenters noted that the constraints EPA placed on spatial averaging proposed in 40 CFR Part 58 would be inadequate to prevent manipulation of spatial averaging to avoid pollution abatement. Several commenters raised environmental justice concerns, noting that minority communities or low income communities might live disproportionately in areas with higher concentrations.

Response: See preamble to the final rule, section II-E-1.

EPA agrees with commenters that the form of the standards, in conjunction with other components of the standards, must protect public health adequately against risks associated with PM. It was for this reason that EPA proposed a policy approach providing for maximum risk reduction for citizens in the community from exposures to the mix of urban and regional scale PM pollution most strongly associated with health effects. In specifically considering whether to allow for the use of spatial averaging, EPA placed great weight on consistency with the underlying body of health effects evidence. EPA notes that some of the commenters opposed to spatial averaging may not have fully understood the implications of the specific constraints and siting requirements discussed in the proposed revisions to 40 CFR Part 58. These constraints, summarized below, are intended to ensure that designated monitors will be reflective of community-wide exposures and that spatial averaging will not encourage "gaming" to avoid control through inclusion of non-representative monitored values from either "clean areas" or "dirty areas."

EPA is concerned that for some commenters, the term "population-oriented monitors," which was used in the proposal to indicate the kinds of sites to be included in spatial averaging, may have conveyed the impression that EPA was focused on protecting only high population centers or locations, rather than smaller communities, whether within urbanized areas or not. This term, derived from air pollution monitoring guidance, simply means a site intended to measure population exposures, as opposed to maximum source impacts or transport. To clarify the intent, EPA has changed the term as it applies to monitors used in spatial averaging to "community-oriented monitors" in the final rules and guidance. The final rule also clarifies that either a single properly sited community-oriented monitor or an average of several such monitors may be appropriate indices of area-wide population exposures for the annual standard, and that both are consistent with

monitoring approaches used in the community epidemiological studies upon which the standards are based.

In EPA's view, the final criteria and siting requirements contained in 40 CFR Part 58 address commenters' concerns regarding the adequacy of the proposed constraints on spatial averaging. The final requirements include provisions for a homogeneity constraint of +/- 20%. This limit governs which monitors may be averaged together such that no monitor may be more than 20% higher or lower than the average of all monitors. The final rules also include provisions requiring that state monitoring plans be available for public inspection. Additional requirements include demonstrations that the monitors to be averaged are influenced primarily by similar sources (e.g., to prevent the placement of monitors upwind in unrepresentative locations), EPA oversight of the monitoring program, which includes regular review and approval of state monitoring plans, and other criteria to ensure proper monitor siting.

Accordingly, the Agency continues to believe that an annual $PM_{2.5}$ standard reflective of area-wide exposures, in conjunction with a 24-hour standard designed to provide adequate protection against localized peak or seasonal $PM_{2.5}$ levels, represents the most appropriate approach to protection of public health against the effects of PM reported in the scientific literature.

(2) *Comment:* Some commenters agreed that spatial averages are closely related to the underlying air quality data used in the health studies and advocated the extension of spatial averaging, both to broad urban scales, as well as to the daily form of the standard; some also recommended less constrained spatial averaging to allow for averaging across entire metropolitan areas.

The Administrator is mindful that some community studies relied inherently on exposure and effects estimates that reflect comparatively broad spatial scales, as highlighted by those commenters desiring to extend permissible averaging. For example, the daily mortality studies generally use urban or metro-areawide effects statistics in conjunction with single or multiple monitors that index day-to-day pollution changes across the area. Ito et al. (1995) found that spatial averages from multiple PM_{10} monitors in Chicago were better correlated with daily mortality than were most single monitors, but that single monitors were also associated. A number of morbidity studies (e.g., Schwartz, 1994; Neas et al., 1995; Raizenne et al., 1996), however, used community-oriented monitors and effects information from a defined group of subjects drawn from the community, who were more closely represented by the monitor.

The combination of studies suggests that extension of spatial averaging to even broader scales than permitted in the proposal would go beyond the exposure regimes of some community-oriented studies. Moreover, as indicated by a number of commenters opposed to any averaging or supporting tighter constraints, such an extension could result in encouraging the inappropriate manipulation of the averages to avoid controls. EPA

believes unconstrained averaging would not be appropriate for all circumstances and might leave some areas without adequate protection. Furthermore, because the 24-hour standard is designed explicitly to address localized peaks, it would be inappropriate to extend spatial averaging to this standard.

(3) *Comment:* One commenter (API, IV-D-2247) suggested that the spatial averaging area should relate to the area that a typical individual might cover during a day to more accurately represent such an individual.

Response: EPA agrees that exposure considerations are important to selecting forms of the standards, but notes that the concentration-response information on PM is derived not from individual exposure studies, but from community population studies. These relationships already factor in population movement over the course of a year. EPA believes that available information in the Criteria Document, as well as prudent health policy considerations, more readily support the spatial averaging approach adopted in the final rule.

(4) *Comment:* The majority of comments from States stressed the need for flexibility in specifying network designs and spatial averaging given that the nature and sources of particle pollution vary from one area to another. One State agency specifically requested the flexibility to choose whether to use a single community-based monitor or a spatial average of several such monitors, arguing that it is appropriate to provide this flexibility as $PM_{2.5}$ monitoring networks evolve and to address the diversity of local conditions. Several comments from States suggested that a spatially averaged form would be difficult to communicate to the general public. Many comments from States stressed the need for additional funding for implementation of spatial averaging in new fine particle monitoring networks.

Response: See preamble to the final rule, section II-E-1.

As noted above, in response to these and other comments, the requirements contained in Appendix N and 40 CFR Part 58 have been revised to clarify that the implementing agencies have the flexibility to compare the annual $PM_{2.5}$ standard either to the measured value at a single representative community-oriented monitoring site, or to the value resulting from an average of community-oriented monitoring sites that meet the criteria and constraints enumerated in the 40 FR Part 58 notice.

EPA agrees with the importance of communicating information about air quality and risk to the public. EPA continues to believe, however, that an annual $PM_{2.5}$ standard reflective of area-wide exposures, in conjunction with a 24-hour standard designed to provide adequate protection against localized peak or seasonal $PM_{2.5}$ levels, is the most appropriate approach to protect public health against the effects of PM reported in the scientific literature. EPA will continue to work with the States and local governments to communicate this information in the context of the fine particle standards.

The issue of funding for monitoring networks raised by some commenters is not germane to the decision on standard form and is addressed elsewhere.

ii. Form of 24-hour standard

(1) *Comment:* A number of commenters maintained that EPA should use a 1-expected-exceedance form for the 24-hour $PM_{2.5}$ standard to limit the number of days per year that the standard is exceeded, as opposed to a concentration based form.

Response: See preamble to the final rule, section II-E-2.

As discussed in the Staff Paper, the proposal, and preamble to the final rule, since promulgation of the current 24-hour PM_{10} standard in 1987, a number of concerns have been raised about the 1-expected-exceedance form. These include, in particular, the year-to-year stability of the number of exceedances, the stability of the attainment status of an area, and the complex data handling conventions specified in Appendix N, including the procedures for making adjustments for missing data and less-than-every-day monitoring. In light of these concerns, the Staff Paper and several CASAC panel members (Wolff, 1996b) recommended that consideration be given to adoption of a more stable and robust form for 24-hour standards.

These commenters apparently gave little consideration to EPA's rationale that a concentration-based form is more directly related to ambient PM concentrations that are associated with health effects because it takes into account the magnitude of PM concentrations, not just whether the concentrations are above a specific level. These commenters also discounted the other advantages of a concentration-based percentile form outlined in the preamble to the final rule. Many of these commenters argue that a 1-expected-exceedance form offers an inherently more stringent level of protection. The level of protection is, however, a function of not only the form, but the level and averaging time. In this case, the 24-hour standard is intended to provide a supplement to protect against 24-hour peaks, thereby supplementing the protection provided by the annual standard with regard to both annual and 24-hour exposures.

EPA continues to believe that a concentration-based percentile form is more reflective of the health risk posed by peak PM concentrations, because it gives proportionally greater weight to days when concentrations are well above the level of the standard than to days when the concentrations are just above the standard. This factor, coupled with the other advantages outlined above, leads EPA to conclude that a concentration-based percentile form will provide for more effective health protection than a 1-expected-exceedance form. EPA believes that the final combination of level and form of the 24-hour standard, in conjunction with the final annual standard, will protect the public health with an adequate margin of safety.

(2) *Comment:* A substantial number of commenters who either supported $PM_{2.5}$ standards

(e.g., Great Basin Unified APCD, IV-D-6502) or provided "contingent" comments if such standards were adopted, strongly supported EPA's move away from the use of extreme values in attainment and planning decisions. Other commenters (e.g., ALA IV-D-2342) expressed the view that stability should not be a consideration, only health protectiveness.

Response: EPA agrees with commenters who suggest that standards can provide more stable targets, while at the same time providing adequate public health protection. EPA generally agrees that increased stability in the standards is desirable to avoid areas "flip-flopping" in and out of attainment, but disagrees that this goal is unrelated to health. In EPA's view, providing a more stable target for controls and more certainty in planning ultimately translates into improved health protection. See preamble to the final rule, section II-B and II-E-2.

(3) *Comment*: Many commenters supporting a single exceedance form or a more restrictive concentration-based percentile form (e.g. a 99thpercentile) expressed concern that the proposed 98th percentile could allow too many high concentration excursions, and thus fail to provide adequate protection against localized peaks, or seasonal emissions problems. Specifically, some commenters expressed theoretically based concerns that in areas with strongly seasonal emissions, such as western areas with winter inversions, over a three year period an area could experience several excursions in which levels could reach as high as 250 $\mu g/m^3$ and still comply with both the annual and daily standards if the remainder of the days had low levels (e.g., 10 $\mu g/m^3$). Others objected to the proposal to limit attainment determinations to population-oriented monitoring sites, and expressed concerns that this would not adequately protect individuals or small communities located in or near localized "hot spots" with high PM levels.

Response: While commenters provided information on peak to mean ratios suggesting the theoretical possibility of such high $PM_{2.5}$ concentrations, there is little evidence that levels this high have occurred in recent years in U.S. urban areas, and the occurrence of such levels under the new standards is even less likely. To address the rare circumstances that might lead to such episodes, EPA intends to establish a significant harm level for $PM_{2.5}$ and associated guidance so States can develop appropriate emergency episode plans. The significant harm and episode criteria will be included in forthcoming guidance. In the interim, the existing PM_{10} emergency episode plans should be triggered by events of the magnitude suggested by these commenters.

EPA shares the concerns of these commenters that the 24-hour standard protect against localized peaks. As detailed in section II.D of the preamble to the final rule, one of the major purposes of the 24-hour standard is to supplement the overall protection of the annual standard in areas with periodic high peak values, including zones of high concentrations due to local sources. To provide adequate safeguards for those who live in or otherwise encounter such localized hot spot areas, attainment of the final standards will be determined by comparison with any population-oriented site within an area. As

discussed specifically in the Part 58 final rule, population-oriented monitoring includes sites in residential areas, commercial areas, recreational areas, industrial areas, and other areas where a substantial number of people may spend a significant fraction of their day. In addition, EPA is modifying the Part 58 requirements to increase the required frequency of sampling in areas where the 24-hour standard is likely to be the controlling standard. This modification will more effectively detect infrequent peak concentrations in such areas.

EPA believes that, given the limits on truly episodic peak concentrations, an appropriately selected 24-hour standard with a concentration based 98th percentile form can provide a stable and adequately protective supplement to the annual standard in areas with periodic peak concentrations.

(4) *Comment:* Other commenters, who opposed setting $PM_{2.5}$ standards, recommended that if EPA proceeds with standards, a concentration percentile form should be used. Further, some such commenters recommended that alternative lower percentiles (e.g., 95th percentiles) be used to provide even more stability.

Response: As discussed in the proposal notice, EPA continues to hold the view that a 90^{th} to 95^{th} percentile form would not serve as an effective supplement to the annual standard because these forms would allow a large number of days with peak $PM_{2.5}$ concentrations above the level of the standard. Specifically, the 90^{th} and 95^{th} percentiles correspond to the 37th and 19th highest daily values in a 365-day data base, respectively.

(5) *Comment:* Several comments from States suggested that a 98^{th} percentile form would be difficult to communicate to the general public.

Response: EPA continues to believe that a concentration-based form of the 24-hour $PM_{2.5}$ standard designed to provide supplemental protection against localized peak or seasonal $PM_{2.5}$ levels, in conjunction with the annual standard, represents the most appropriate approach for protection of public health against health effects from exposure to PM. Although the statistical form of the standard may be a complex subject, EPA believes that the advantages of the form can and should be explained to the interested public. Concepts such as percentiles are routinely presented to the general public in other venues (e.g., interpreting students' scores on standardized tests, infants' height and weight statistics). As described above, EPA agrees with the importance of communicating information about air quality and risk. EPA will continue to work with the States and local governments to communicate this information in the context of the fine particle standards.

(6) *Comment:* Some commenters recommended that EPA should adopt a form of the standard using a statistical test that creates a "too-close-to-call" category to reduce attainment flip flops.

Response: EPA has considered and rejected this approach for achieving stability in the form of the standard. In essence, if controls are applied to such a category, it is effectively non-attainment. If no controls are applied, then the "too-close-to-call" category effectively becomes an attainment category and the "level" of the standard has effectively been raised. While the use of a statistical test can reduce the misclassification rate (i.e., cases where truly attainment areas are classified as nonattainment, and vice versa), it can also delay the implementation of needed controls or, conversely, the time for an area to be redesignated from nonattainment back to attainment. The revised PM standards are an annual mean (which is very stable) and a concentration-based form of the 24-hour standard that is more stable than the one-exceedance standard used in the past.

(7) *Comment*: A number of commenters disagreed with averaging over three years for both the annual and 24-hour standards because of their desire for quick action in the initial implementation of $PM_{2.5}$ controls.

Response: The EPA recognizes the importance of promptly implementing appropriate control programs, but EPA does not believe that implementation start-up concerns are an adequate basis for adopting a form (e.g., a single year annual average) that would provide less stable risk reduction in the long-run. Therefore, the EPA continues to concur with the Staff Paper recommendation, supported by CASAC, to use the annual arithmetic mean and 98th percentile, averaged over 3 years, as the forms of the $PM_{2.5}$ standards. Nevertheless, EPA intends to address the concerns of those who commented that the 3-year form might prevent the public from being informed about the air quality status of their communities. As outlined in section II.H of the pramble to the final rule, EPA plans to issue revised Pollutant Standard Index criteria for $PM_{2.5}$, to ensure the public is informed promptly about air quality status.

e. Revisions to Appendix K for PM standards

The final rule for PM retains Appendix K to 40 CFR part 50 in its current form for use in connection with the continued applicability of the current PM_{10} standards during a limited transition period (see section VII of the preamble to the final rule). EPA's final action on the proposed revisions to Appendix K, intended for use with the revised PM NAAQS, is incorporated in a revised versionj of Appendix K that is designated as Appendix N in the final rule. Because the comments and issues related to the proposed revisions to Appendix K (now Appendix N) are similar for both the PM_{10} and $PM_{2.5}$ standards, this section of the response to comments addresses comments included in both section II.A.2.a.(5) and section II.A.2.b.(5) in the Summary of Comments. For clarity, the statements of comments and EPA responses below use the designation Appendix N in referring to the proposed and final revisions to Appendix K.

i. <u>Data completeness and missing data adjustments</u>

(1) *Comment:* Some commenters disagree with retaining high 24-hour values when the 75%

data completeness requirement is not met, yet not retaining low values when the 75% data completeness requirement is not met.

Response: The EPA response to this comment is addressed in section V of the preamble to the final rule.

(2) *Specific Comment* (State of North Carolina Office of the Governor, IV-D-7003): The commenter specifically questioned how the requirement for having 3 consecutive years of air quality data is affected by a situation in which years 1 and 3 have data above the standard while year 2 has incomplete data with no values above the standard.

Response: The EPA agrees that in some cases it makes sense to retain a particular year and use the data to show that the standard has or has not been met even though the year may have one or more quarters with less than complete data. Appendix N allows some cases of less than complete data to be used subject to the approval of the appropriate Regional Administrator. Thus, the requirement for 3 consecutive, complete years of air quality data still exists, but exceptions to having complete data may be allowed. The EPA plans to issue guidance on dealing with less than complete data.

(3) *Specific Comment* (State of North Carolina Office of the Governor, IV-D-7003): The commenter stated that to comply with the calculation of the 98th percentile a monitor must have at least 49 samples in a calendar year.

Response: Although there are many methods to define or calculate a percentile from a discrete set of data, the EPA has selected a method that is simple, is not biased high or low for all sample sizes, and yields reasonable results. The method as proposed and incorporated into Appendix N does not limit the sample size to above 49. The method used by the EPA results in the maximum value from 1 to 50 samples for a 98th percentile (1 to 100 samples for a 99th percentile). Therefore, the form of the standard does not limit the sample size. However, the EPA has set various data completeness requirements that must be satisfied when using data to compare with the levels of the standards.

(4) *Comment:* A number of air pollution control agencies were concerned that Appendix N does not allow seasonal sampling. They argue in some areas it is appropriate to conduct seasonal sampling, reducing the frequency of monitoring during a period of expected low concentrations to save resources.

Response: Appendix N does not prohibit seasonal sampling, and refers matters of sampling frequency to 40 CFR Part 58.13. For clarification, the final Appendix N adds that exceptions to specified sampling frequencies, such as a reduced frequency during a season of expected low concentrations, shall be subject to the approval of the appropriate Regional Administrator. The EPA plans to issue guidance pertaining to exceptions to sampling frequency.

(5) *Comment:* Appendix N should address the data affected by special or uncontrollable events.

Response: Several commenters mentioned the handling of special or uncontrollable events in conjunction with Appendix N. The EPA believes that data resulting from uncontrollable or natural events, for example, fires or high winds, may require special consideration. In some cases, it may be appropriate to exclude these data because they could result in inappropriate values to compare with the levels of the PM standards. In other cases, it may be more appropriate to retain the data for comparison with the level of the PM standards and then allow the EPA to formulate the appropriate regulatory response. Appendix N clarifies that exclusion, retention, or adjustment of the data affected by uncontrollable or natural events is subject to the approval of the appropriate Regional Administrator. All relevant EPA guidance should be considered.

ii. Data handling and rounding conventions

(1) *Comment:* The requirement to compute quarterly averages when calculating the annual mean is unnecessary and too cumbersome to justify.

Response: The EPA considered not using quarterly means to compute annual means; however, the EPA disagrees that there is no justification for using quarterly means. Significant seasonal variation is observed in PM concentrations. Seasonal variation has also been observed in data completeness and sampling schedules. Computing an annual mean with quarters sampled at different rates could lead to an annual mean that is not representative of PM levels in an area. Calculating quarterly means with which to compute annual averages avoids this bias. The EPA, therefore, has retained the use of quarterly means when calculating an annual mean.

(2) *Comment:* The mathematical formulas for computing the percentiles are incorrect. Picking the next higher value or averaging the next highest value make the percentile too high. The formulas for computing the percentiles are too complex and are difficult to understand and explain to the public. The EPA formulas for 98th percentile are not defined for less than 48 samples in a year.

Response: EPA believes that the computational formulas in Appendix N are mathematically correct and consistent with the definition of a percentile. The convention of using the next highest value insures that the specified percentage of data is below the computed value. In Appendix N, EPA has altered the calculation from the proposed formula in an effort to simplify the calculation. Further, the EPA believes that a percentile form of the standards is easier to explain than an exceedance form which requires complex adjustments for different sampling schedules. The EPA also disagrees that the formula for the 98th percentile is not defined for less than 48 samples in a year. In fact, the formula for the 98th percentile will yield the maximum value in a year until the sample size is greater than 50, and the formula for the 99th percentile (in the case of

PM_{10}) will yield the maximum value until the sample size is greater than 100.

(3) *Comment:* The annual and 24-hour forms of PM_{10} and $PM_{2.5}$ standards should have the same decimal precision. These differences are inconsistent with quantitative uncertainty and confusing.

Response: The EPA realizes that the different levels of precision for the standards can be confusing. The EPA disagrees that this is inconsistent with measurement precision and quantitative precision. The different levels of precision used for the annual and the 24-hour forms of the PM standards are due to the different levels of information in each form of the standard. The annual form is a mean involving all values in 3 years while the 24-hour is based on only 3 values (the 98th or 99th percentile from each year). The statistical characteristics of these two quantities are completely different. To regard them as having the same level of precision is unjustified. Therefore, the EPA has retained the different levels of precision in the annual and 24-hour forms of the PM standards.

(4) *Comment:* The rounding conventions used for the annual and 24-hour forms of the PM standards should be the same. Rounding up or down should not be allowed before comparing to the level of the standards.

Response: The EPA considered the complexity of different rounding conventions for the annual and 24-hour formulas of the PM standards. However, the rounding conventions proposed by the EPA reflect the precision inherent in the calculations for the different forms of the standard and mechanically follow standard mathematical practice of rounding up at 5 or greater and down at 4 or less (e.g. rounding 65.5 to the nearest 1 would be 66 while 65.4 would be 65). The EPA also feels that rounding is necessary before comparing to the level of the standard. Rounding is justified due to the actual measurement precision experienced in measuring PM. To ignore measurement precision is not scientifically defensible. Therefore, the EPA has retained the proposed rounding conventions in the revised Appendix N.

iii. 3-Year compliance period

(1) *Comment:* The compliance test for the annual form of the PM standards should be on the basis of a rolling average of 12 quarters instead of 3 years. This would allow measured data to demonstrate compliance as much as a year earlier than the 3 year rolling average.

Response: The EPA agrees that compliance could be demonstrated earlier in some cases with measured data on the basis of a 12 quarter rolling average. However, adopting this form of the annual standard is not justified since all other PM standards will be on a 3 calendar year basis. Different periods for compliance demonstrations could lead to unnecessary confusion as to the actual status of an area and confusion as to what control strategies should be employed and when. The EPA has retained the 3 calendar year basis for both $PM_{2.5}$ and PM_{10} annual forms of the standard.

3. Specific comments on proposed primary PM_{10} standards

a. Indicator for coarse particles

Relatively few public comments were received in this area. Public comments received on the proposed indicators were overwhelmingly in favor of EPA's proposal to maintain PM_{10} as an indicator for PM, whether as an indicator of coarse particles in conjunction with a fine PM standard, or as the sole PM indicator. This near unanimity shows strong support for retaining general PM standards. The comments are summarized primarily in section II.A.2.b.(1) of the Summary of Comments and significant comments are highlighted in the preamble to the final rule.

(1) *Comment:* Several CASAC panel members suggested using the coarse fraction (i.e., $PM_{10-2.5}$) directly as the indicator.

Response: See preamble to the final rule, sections II-C-2.

(2) Comment: One commenter (NMA, IV-D-2158), although supporting the PM_{10} indicator, suggested that the coarse fraction be subtracted from the PM_{10} mass measured.

Response: A literal reading of this suggestion shows it would amount to establishing a standard for fine particles without a PM_{10} standard. These commenters base their recommendation on their assessment of the scientific evidence, which in their opinion effectively supports EPA's conclusions regarding the desirability of separating fine and coarse particles, but also concludes that coarse fraction particles at current and anticipated levels have no substantial health and welfare effects. EPA believes that establishing an indicator for fine PM is essentially consistent with these commenters' first recommendation, but disagrees that the evidence in any way warrants total elimination of coarse fraction PM standards and controls.

b. Averaging times

The comments addressed in this section are summarized primarily in section II.A.2.b.(2) of the Summary of Comments and significant comments are highlighted in the preamble to the final rule. Most comments focused on the proposed alternative of revoking the 24-hour PM_{10} standard. EPA received few comments supporting elimination of the 24-hour PM_{10} standard. The main exception were some industries, most notably the mining industry, who argued that the available data provide little evidence for coarse particle effects at current ambient levels. These groups, who generally opposed $PM_{2.5}$ standards, also argued that the daily PM_{10} standard could be eliminated if $PM_{2.5}$ standards were set. EPA has responded in full to these comments in sections II.B and II.G.1 of the preamble to the final rule.

c. Standard levels

The comments addressed in this section are summarized primarily in section II.A.2.b.(3) of the Summary of Comments. Significant comments are highlighted in sections II.B and II.G.2 of the preamble to the final rule.

(1) *Comment:* Colorado PIRG & Colorado ALA (IV-D-2095), asserted that EPA has "de-emphasized potential effects of coarse particles even though the available epidemiological evidence does not imply that acute exposure to coarse particles is benign or without effect."

Response: EPA does not agree that the effects of coarse particles are de-emphasized in the current rulemaking procedure. In its decision to retain standards for PM_{10}, EPA noted that certain health effects are plausibly associated with both long- and short-term exposures to coarse fraction particles. The studies cited by the commenter (Gordian et al., 1996; Hefflin et al., 1994) are discussed in the conclusions to the Staff Paper (p. VII-37) as evidence of these health effects. Both coarse particle studies involved unusually high exposures to fugitive dust or airborne volcanic ash, and both measured some increases in respiratory illness (asthma or bronchitis) with increases in PM_{10} that were primarily coarse particles.

As summarized in the Criteria Document (pp. 12-373 to12-377), where measures of both fine and coarse particles were used in health effects investigations, the researchers often found stronger associations with the fine particles or components. For example, in the Harvard Six Cities study (Dockery et al., 1993), of the three PM indices, $PM_{2.5}$ was most predictive of mortality for five of the six cities, but there were still associations found with coarse particle measures. EPA believes that the available scientific information indicates the need for additional regulatory attention to the fine component of PM. Although EPA has modified the form of the 24-hour PM_{10} standard to account for the additional protection provided by the additional $PM_{2.5}$ standards, both 24-hour and annual PM_{10} levels have been retained to provide adequate protection against the known and potential effects associated with the larger inhalable coarse fraction particles.

(2) *Comment:* The Ontario Ministry of Environment and Energy (IV-D-5262) urged the EPA to set more stringent standards for PM_{10}, stating: "The recent study in Anchorage, Alaska shows adverse effect associations with the coarse fraction. The study is discussed in the Health Effects section of the staff paper, but does not appear to have been considered in the decision making." In addition, the Clean Air Coalition (IV-D-7730) urged the EPA to "adopt California's PM_{10} standard which is two thirds stronger than the national standard." The commenters attached a copy of the recent scientific article by Gordian et al. (1996) in support of their comments. By contrast, the Alaska Department of Health and Social Services, Division of Public Health (IV-D-7005), in supporting EPA's proposed PM_{10} standards, offered a critical review of a recent report from the Gordian et al. (1996) study and argued that it should not be relied upon in the development of PM standards.

Response: EPA does not agree that the results of this study were not considered in the decision-making process; the Anchorage study was among the many studies whose findings were used in determining appropriate levels of PM standards. In particular, this study (Gordian et al., 1996) is one of the few that found evidence of health effects associations with PM_{10} that was conducted in an area where PM mass appears to be dominated by coarse fraction particles. As commenters have observed, the results of this study were reviewed and considered in the context of a limited set of other studies of exposure to PM_{10} with relatively high coarse fraction concentrations.

As noted in the Staff Paper, the 22-month average PM_{10} concentration in this study of 45.5 $\mu g/m^3$ is near the level of the annual average PM_{10} standard, and over the course of the study, multiple exceedances of the 24-hour standard occurred, with extreme values during the volcanic activity that occurred near the beginning of the study. In examining this study in the context of other relevant PM_{10} studies, staff noted some methodological weaknesses, but generally found the results supported retention of PM_{10} standards at levels at or near those of the current standards. Based on the overall evaluation of the scientific evidence, EPA believes that it is appropriate to maintain the current annual average standard for PM_{10} at 50 $\mu g/m^3$, and to retain the level of the 24-hour standard, but with a more robust concentration based form based on the 3-year average of the 99^{th} percentile 24-hour values.

(3) *Comment:* Some commenters, including some environmental groups and the State of California (Cal EPA, IV-D-2251), suggested that the large number of recent studies showing effects at PM_{10} levels below the current standards provides a basis for establishing stricter annual and 24-hour PM_{10} standards, in conjunction with $PM_{2.5}$ standards.

Response: See preamble to the final rule, sections II.B and II.G.2.

As discussed in Sections II.B and C, while these studies could be interpreted as supporting either a tightening of the PM_{10} standards or the addition of standards that focus control on the fine fraction of PM_{10}, the weight of evidence from all of the relevant information more readily supports the development of additional protection for the $PM_{2.5}$ fraction. Given EPA's policy approach of useing PM_{10} as a coarse particle indicator, in EPA's judgment it is not appropriate to use these studies for both the $PM_{2.5}$ and PM_{10} standards.

d. Form of 24-hour standard

In general, comments received on the form of the 24-hour PM_{10} standard paralleled those on the form of the $PM_{2.5}$ standard. Substantial concerns were expressed by environmental groups, some states, and others that the 98^{th} percentile would not provide an adequate limit on the number and magnitude of 24-hour peak PM_{10} excursions. While a number of these commenters suggested keeping the current 1-expected-exceedance form, EPA believes that a

concentration based percentile form offers significant advantages, as discussed in section II.G of the preamble to the final rule and in the responses above in section 2.d, for both PM indicators.

Comment: Many commenters were concerned that the uncertainties in the available scientific information on the effects of coarse particles were a reason to be concerned that, assuming the current standard level was kept, a 98th percentile form would represent a significant relaxation in protection relative to the current standards. Some of these commenters recommended keeping the current 1 expected exceedance form, while others argued that the current form and level be maintained during a transition period, to prevent inappropriate reduction of health protections in interm implementation. Some air pollution control officials suggested consideration of a 99th percentile form with increased monitoring as an appropriately protective form.

Response: See section II. G of the preamble to the final rule. EPA notes that, unlike the situation for the new $PM_{2.5}$ standards, in the case of the PM_{10} standards, the 24-hour standard has generally been the "controlling" standard, making changes to the form of the 24-hour standard potentially more significant to the overall national level of protection afforded. It is clear that, if the level of the current standard is held constant, as suggested by staff and a number of CASAC panelists, the move to a 99th percentile could result in relaxation of current controls on coarse particles. Given the uncertainties with respect to the potential health effects of coarse fraction particles, the Administrator is persuaded that the somewhat more cautious approach recommended by many commenters is appropriate. For these reasons, EPA modified the final rule to express the daily standard in a 99th percentile concentration-based form, averaged over 3 years.

EPA shares commenters' concern with respect to keeping appropriate protection during the transition to the revised and new PM standards. Accordingly, as discussed in section VII of the preamble to the final rule, the current PM_{10} NAAQS will continue to apply during the transition period.

e. Revisions to Appendix K for PM_{10}

See combined response to comments on proposed revisions to appendix K (appendix N in the final rule) in section II.A.2.e above, as well as section V of the preamble to the final rule.

4. Specific scientific/technical comments

a. Interpretation of epidemiological studies

This section addresses comments included in section II of the Summary of Comments..

i) Comments on Specific Studies

(1) *Comment:* Many commenters (e.g., UARG IV-D-2250) asserted that EPA was selective

in its use of epidemiological evidence, and cites as an "extreme example" the examination of evidence of pollutant effects on respiratory function in children. In particular, the commenter argues that EPA ignored the following conclusions drawn from the 24-city study by Dockery et al. (1996): "There appears to be *no evidence* that the prevalence of asthma or asthmatic symptoms in children is associated with chronic exposure to particulate, sulfur oxide, or ozone air pollution ... [A]ir pollution *does not appear to contribute* to the increased prevalence of new cases of asthma, as is often claimed in the popular press." (UARG p. 21, emphasis added by commenter). Commenter notes this study is not included in the summary of studies that appear in Figure 1 of the proposal notice.

Response: A general response to the comment that EPA ignored or downplayed numerous studies is contained in Section II.B of the preamble to the final rule, and in Appendix B of this document. The commenter's specific criticism that EPA did not include Dockery et al. (1996) in Figure 1 of the proposal preamble is misplaced. That figure includes only studies that were considered to represent short-term exposure studies, and the 24-city study is clearly a long-term exposure study. It is important to note that EPA outlined key information from the Staff Paper and the Criteria Document in the preamble to the proposed rule, but did not attempt to repeat in full the complete analysis of the literature that was summarized in the Criteria Document.

The commenter also asserted that the EPA has taken a biased, or selective approach in its review of the scientific literature. EPA accurately reported the findings of Dockery et al. (1996) with respect to the study not finding significant associations between pollutant exposures and all health endpoints measured. In fact, the specific observation made by the commenter -- that there were no statistically significant associations found between the pollutants and asthma symptoms or wheeze -- was noted in the Criteria Document (p. 12-187).

In their discussion of this study's findings, Dockery et al. (1996) distinguish between findings of asthma symptoms and bronchitic symptoms. The authors cite previous findings of earlier studies as supporting their observation that exposure to particles is not associated with increased reporting of asthma symptoms over an annual period. In the text as it is excerpted by UARG, it appears that Dockery and colleagues are arguing against potential respiratory health effects from PM. In fact, however, the results of this study support the original hypothesis that acid aerosol exposures can affect childrens' respiratory health. In their discussion, the authors conclude only that air pollution exposure does not appear to be associated with the <u>development</u> of asthma or wheeze, based on the lack of significant findings for asthma or wheezing. The full quote from Dockery et al. (1996) clearly indicates this:

> "There appears to be no evidence that the prevalence of asthma or asthmatic symptoms in children is associated with chronic exposure to particulate, sulfur oxide, or ozone air pollution. This does not imply that

> acute air pollution episodes do not trigger or exacerbate asthmatic attacks, as has been shown in multiple studies. However, air pollution does not appear to contribute to the increased prevalence of new cases of asthma, as is often claimed in the popular press."

Indeed, a number of other commenters on the proposed PM standards, as well as the proposed ozone standards, have erroneously characterized EPA as claiming that the observed increase in asthma prevalence is demonstrably related to air pollution, or that the observed the increase in asthma is not coherent with the decline in air pollution. As noted elsewhere, EPA has not made such claims, and for the reasons outlined by Dockery et al. (1996) above, does not believe there is any lack of coherence with respect to these observed trends and the observation from short-term exposure studies that acute air pollution excursions can trigger or aggravate asthma symptoms.

As pointed out in the PM Criteria Document, both short-term and long-term exposure epidemiological studies find that ambient PM is associated with decreased pulmonary function and increases in lower respiratory symptoms. In the companion report from the 24-city study, Raizenne et al. (1996) found significant decreases in several lung function parameters (FVC, $FEV_{1.0}$, $FEV_{0.75}$) with all measures of particulate matter. It may or may not be true that exposure to air pollutants plays some role in the increased prevalence of asthma found in many countries. Of clear concern here are not only the instant observation of increased symptoms and decreased lung function, but the potential consequences of longer-duration exposure in later life. As stated by Dockery and colleagues in their concluding paragraph: "These data indicate that chronic exposure to acid aerosol pollution may have observable negative consequences for the health of children. Although the long-term consequences of bronchitis in these children remain unclear, respiratory illnesses in childhood may be a risk factor for chronic obstructive disease..."

(2) *Comment:* Lipfert (IV-D-2180) also criticizes EPA's interpretation of the 24-city studies (Dockery et al., 1996; Raizenne et al., 1996), asserting that the study fails to show effects within the high-sulfates regions, but only shows effects between this region and others.

Response: The commenter focuses on the lung function finding, apparently placing greater weight on the graphical depiction of the results than on the accompanying statistical analyses of the data. The authors (Raizenne et al., 1996) report a statistically significant association between particle strong acidity and decreased FVC for the subgroup of Eastern cities where there are generally higher levels of acid aerosols and sulfates. These negative associations are found in all subregional analyses, and the authors conclude that "...no differences in the association were observed with regard to geography..." This would appear to contradict the commenters' observation. Even if the comment were correct, however, this finding would suggest a relation between acid sulfates and the observed effects. Moreover the 24-city studies have several features that serve to greatly reduce problems associated with traditional cross-sectional mortality

studies of air pollution: (1) the quality and extent of monitoring in the study are high, (2) the suburban locations and young age of the subjects preclude significant concerns about much higher historical concentrations, and (3) the extent of subject-specific information on lifestyle and other potentially important co-factors in this cohort study. These features increase confidence in the results of the study.

(3) *Comment:* AISI (IV-D-2242), in asserting selective use of scientific studies by EPA, cites the use of Saldiva et al. (1995) but not Saldiva et al. (1994) in the Staff Paper.

Response: Both papers were reviewed in the Criteria Document and contained in the 38 studies in Table 12-2, all of which were cited by reference to the Criteria Document table in the short-term exposure mortality discussion in the Staff Paper. Saldiva et al. (1995) is not given any significant discussion in the Staff Paper, although it is referenced. The Staff Paper does not and cannot completely replicate the analysis from the Criteria Document, but focuses on key studies highlighted in the Criteria Document. Thus, EPA does not agree that it has overlooked the results of Saldiva et al. (1994) in its review of the PM-related health effects literature.

As stated in the Criteria Document (p. 12-74), Saldiva et al. (1995) found a significant association between PM_{10} and mortality in elderly (>65 years of age) people. In a multi-pollutant model, the association with PM_{10} remained significant but the associations with co-pollutants (SO_2, NO_x, CO) did not. The same data set was used in the earlier study (Saldiva et al., 1994) but the effect under study was mortality for respiratory causes in children under five years of age (Criteria Document, p. 12-76). Among five pollutants considered both individually and in a multipollutant model (CO, SO_2, NO_x, O_3 and PM_{10}), only an association with NO_x was found to be statistically significant. In Table V-9 of the Staff Paper, EPA summarized the Criteria Document assessment of subpopulations especially sensitive to PM-related health effects; the elderly are considered to be a sensitive subpopulation with regard to mortality from acute exposure to PM, but no conclusions are drawn ("insufficient data" or +/-) for PM-related mortality in children. This is clearly consistent with the then available literature, including both Saldiva studies. EPA therefore does not agree with the commenter's assertion that the literature was used selectively. EPA believes that the two papers by Saldiva and colleagues were adequately characterized in the Criteria Document and that the Staff Paper accurately summarized the salient results.

(4) *Comment:* Several commenters assert that the EPA has relied excessively on a few individual scientists in reaching conclusions regarding PM and health effects. Engine Manufacturer's Association (IV-D-2328) states: "EPA relies disproportionately upon the opinions expressed by a small but prolific community of researchers (primarily Schwartz and his co-authors, including Pope and Dockery) in reaching its conclusions. Nearly two-thirds (19 out of 29) of the peer-reviewed journal articles cited by the EPA to support its conclusions and interpretations are authored or co-authored by these three investigators." Other commenters made statements that were similar, though differing

somewhat in detail. For example, Exxon Chemical Americas (IV-D-2113) states: "of the 18 studies used as key, ... 15 of 18 involve only three authors (Schwartz, Pope or Thurston)."

Response: EPA does not agree with these commenters' assertions that over-reliance was placed on the findings or opinions of a few scientists. In the first instance, the comments themselves show this is clearly not a situation in which reliance is placed on a single study by a single investigator or group. Instead, a variety of different data sets, modeling procedures, study designs, and approaches have been used in a number of locations around the world. While it is true that Drs. Schwartz, Pope, Dockery and Thurston have conducted numerous studies and published many research reports on pollution-related health effects, an examination of the over 80 key studies summarized in Chapter 12 of the Criteria Document shows that dozens of other scientists have either published their own independent work or collaborated with these investigators on various projects. The fact that particular scientists are prolific or have an active research program does not give cause to question the validity of their research. Moreover, these investigators subjected their results to independent peer review and publication in respected scientific journals. Beyond that, in the development of the Criteria Document, each study was evaluated on its own merits by EPA scientists, outside researchers and the CASAC. A discussion of the consideration of studies by different investigators in the Criteria Document is outlined in Section II.B of the preamble to the final rule and in Appendix B of this Response to Comments.

It is also important to note that a number of the studies by these authors were subjected to reanalyses. In general, the Criteria Document concluded that the most comprehensive and thorough reanalyses were those in the series conducted for the HEI, which reanalyzed data sets used in studies from six urban areas in Phase I.A (Samet et al., 1995). Among others, the reanalyses included studies by Dockery et al. (1992), Pope et al. (1992), Schwartz and Dockery (1992a), and Schwartz (1993). The most important finding in the HEI Phase I.A reanalyses of these studies is "the confirmation of the numerical results of the earlier analyses of all six data sets" (HEI, 1995). After replicating the original investigators' analyses, Samet et al.(1995) also found similar results analyzing the data using an improved statistical model. The HEI Oversight Committee found:

> [I]t is reasonable to conclude that, in these six data sets, daily mortality from all causes combined, and from cardiovascular and respiratory causes in particular, increases as levels of particulate air pollution indexes increase [HEI, 1995].

These reanalyses by respected independent scientists confirm the reliability and reproducibility of the work of the original investigators, which include those most frequently mentioned by the above commenters. Indeed, the level of scrutiny applied to the work of these investigators has given EPA a clearer basis for relying on their work in

the context of the larger body of scientific information on the effects of PM.

(5) *Comment:* Some commenters took issue with EPA's characterization of the results of Thurston et al. (1994), both with respect to suggesting a greater association between hospital admissions and $PM_{2.5}$ than PM_{10}, and with respect to the statistical significance of any PM indicators when ozone was included in the model.

Response: EPA does not agree that the results of Thurston et al. (1994) were mischaracterized in the Staff Paper. The general conclusions drawn by the authors from this study were that ozone exposure "clearly dominates" the summertime air pollution association with hospital admissions, while a "substantial increase in respiratory admissions" is also associated with peak acid aerosol exposures. The results of single-pollutant models in this study show significant associations with all measures of particles but the coarse fraction, and that only H+ remains significant when the particle measures are modeled with ozone. These findings are precisely the results reported in EPA's Staff Paper. The quote from Thurston et al. (1994) cited by UARG -- "although the particulate matter mass measures initially appeared to be significantly correlated with hospital admissions, . . . it seems clear that these apparent associations were merely a statistical by-product of interpollutant confounding results from the shared day-to-day variations in dispersion conditions" -- is drawn from an extensive discussion of the interactions between pollutants in multi-pollutant models. Based on a full reading of the paper, it is clear that the authors are emphasizing the need to address the issue of interpollutant confounding between various measures of PM and other pollutants. The authors find that the associations found for ozone and airborne acids appear to be independent of one another, with the association found between airborne acids and hospital admissions remaining significant when ozone is added to the model. In the area in which the study was conducted, acid aerosols are a large component of airborne particles, and the significant finding of increased hospital admissions with increased airborne acid levels is an indication of an adverse effect of exposure to a major component of airborne particles. In the paper's discussion section, this issue is more fully addressed:

> These analyses considered 10 different pollution variables, of which 7 were various particulate matter metrics. In the initial analysis, virtually every pollutant measure was significantly correlated with respiratory hospital admissions. Even nonthoracic particle mass (TSP-PM_{10}) approached statistical significance, which is biologically implausible. This association is likely due to the moderate to high correlation among the various pollutants, which results from the general influences of atmospheric dispersion conditions on all of them . .. . after controlling for these O_3 associations, the relative particle metric strengths of association with health effects were generally $H^+ > SO_4^= > FP > PM_{10} > TSP$... These results are biologically plausible in that the largest (i.e., nonthoracic) particles are least associated with admissions, while the smallest and most irritant particles (i.e., submicrometer acid aerosols) are indicated to be most strongly associated with adverse effects. [Thurston et al., 1994.]

(6) *Comment:* American Petroleum Institute (IV-D-8045) assserts that there is an extremely limited amount of scientific information on directly measured $PM_{2.5}$. In addition, the commenter asserts that none of the $PM_{2.5}$ air quality data are more recent than 1988, so they have limited, if any, relevance to today's air quality.

Response: EPA disagrees with the commenter's assertions. While there are fewer studies that directly measured fine particles than PM_{10}, there are over 30 epidemiological studies that obtained data on indicators of fine particles. As compared to the data available for the 1987 decision to establish PM_{10} standards that this commenter now supports, far more studies exist today for fine particles than existed at that time for PM_{10}. In the final decision on the levels for the $PM_{2.5}$ standards, the EPA was able to rely 9 studies that used gravimetric measures of fine particles, including $PM_{2.5}$ and the nearly identical $PM_{2.1.}$ Contrary to commenters' assertion, three of these studies (Dockery et al., 1996; Raizenne et al., 1996; Neas et al. 1995) were based on directly measured fine particle data collected in 24 North American cities between 1988 and 1991.

Moreover, the suggestion that North American data more than 10 years old has little relevance to today's air quality has no basis in fact. While PM levels have declined on average, there are numerous examples of studies using data collected in the 1980's that contained PM levels and composition that are well within the ranges that exist today. As discussed in Section II.F of the preamble to the final rule, the long-term average concentrations in a number of studies finding significant fine particle effects is about 18 $\mu g/m^3$. More specifically, two of the six cities with significant associations between mortality and fine particles had annual $PM_{2.5}$ levels of 15.7 $\mu g/m^3$ and 18.6 $\mu g/m3$. The level of the annual standard of 15 $\mu g/m^3$ was selected to be below these levels. $PM_{2.5}$ data from 1993-95 show Washington, D.C. had annual levels of about 18 $\mu g/m^3$, Philadelphia had 17 $\mu g/m^3$, and Los Angeles had 30 $\mu g/m3$. Further, Chapter 6 of the Criteria Document shows that the dominant components of fine particles in various regions of the country have not changed significantly since the 1980's. An examination of the data clearly demonstrates that the commenter is in error and that both the quantity and the composition of the 1980's Six City data is directly relevant to current air quality.[6]

ii) Confounding issues

A large number of commenters expressed concerns that the available epidemiological relied upon by EPA did not adequately take into account a variety of potential confounders, including co-pollutants, weather, personal factors, and indoor air pollution. In particular, some

---

[6]If these commenters are convinced that current levels are, in fact, below the the ranges in the recent studies including Boston and St. Louis and the 24 cities, then their otherwise non-germaine arguments with respect to the high costs of implementing the standards are without merit, on their own terms.

commenters noted particular studies they felt had done a better job at addressing multiple pollutants or other confounders as "negative" studies that EPA ignored and down played. The issue of potential confounding or effects modification by co-occurring pollutants or other factors such as weather in epidemiological studies of PM was central in the review of the scientific criteria. As such, it received substantial attention and analysis in Chapter 12 of the Criteria document, and in Chapters V and elsewhere in the staff paper. The major findings from these documents were also summarized at some length in the proposal. Because this issue has several dimensions, aspects of this issue have been addressed in discussing consistency with respect to certain multi-pollutant studies EPA allegedly ignored (see above and Appendices A, B, and C), weather (see appendices B and C), indoor air pollution and exposure misclassification (see below and Appendix D), as well as in corresponding portions of section II.B. of the preamble to the final rule.

As discussed therein, EPA's assessment of numerous mortality studies concludes that when studies are evaluated on an individual basis, the PM-effects associations are valid and, in a number of studies, not seriously confounded by co-pollutants (U.S. EPA, 1996a; p. 13-57); and when a collection of studies from multiple areas with differing concentrations of PM and co-pollutants are examined together, the association with PM_{10} remains reasonably consistent across a wide range of concentrations of these potentially influential pollutants (U.S. EPA, 1996a, p. 12-33; U.S. EPA, 1996b, p. V-55). EPA has carefully considered the view advanced by some that the results of individual studies of multiple pollutants, such as the HEI Philadelphia studies, are more suggestive of an "air pollution" effect than an effect of PM alone. Indeed, the proposal notes that it is reasonable to expect that other pollutants may play a role in modifying the magnitude of the estimated effects of PM on mortality, either through pollutant interactions or independent effects. Based on the large body of evidence at hand, however, EPA cannot accept the suggestion that such multi-pollutant studies are in any way "negative" with respect to EPA's conclusions that PM, alone or in combination with other pollutants, is associated with adverse effects at levels below those allowed by the current standards. This conclusion is based not only on the consistency of PM effects across areas with widely varying concentrations of potentially confounding copollutants, but also on the extended analyses of the Philadelphia studies in the Criteria Document and Staff Paper.

iii) Statistical modeling issues

Some commenters argued that sufficient flexibility exists in the analyses of large data sets that it may be possible to obtain almost any result desired through choice of statistical method. Analytical choices include the specific statistical model; methods used to adjust for seasonal variation and the trends in the data; treatment of other variables (e.g., other pollutants, weather, day of week); "lag" structure; and study population. The issue of statistical modeling also received substantial attention and analysis in the criteria review, with a substantial discussion in Chapter 12 of the Criteria Document. The Criteria Document concludes that:

> [T]he largely consistent specific results, indicative of significant positive associations of ambient PM exposures and human mortality/morbidity effects, are not model specific,

> nor are they artifactually derived due to misspecification of any specific model. The robustness of the results of different modeling strategies and approaches increases our confidence in their validity [U.S. EPA, 1996a, p. 13-54].

Appendices A, B, and C all discuss statistical modeling issues in the context of particular studies, with a response to the general issue raised here presented in section II of Appendix B, and a summary response in section II.B of the preamble to the final regulation.

iv) Exposure misclassification issues

A range of public comments were received in this area. The comments are summarized primarily in sections II.A.3.a.(4), II.A.3.h., and II.A.3.e. of the Summary of Comments and significant comments are highlighted in the preamble to the final rule. In addition to the responses contained in the preamble to the final rule in section II.B, EPA provides the following additional responses to specific issues. Appendix D contains further technical discussion of exposure-related issues.

(1) *Comment:* The available PM epidemiological studies are flawed because they rely on central monitors, not personal monitors, as an index of exposure to PM. Some commenters further note studies that showed a lack of correlation in cross-sectional comparisons between outdoor PM measured at central locations and indoor or personal exposures to PM (which includes PM from the outdoor, indoor and personal environments).

Response: See section II.B of the preamble to the final rule, and the detailed discussion in Appendix D.

EPA rejects the notion advanced by commenters that epidemiological studies must use personal exposure monitoring to be considered for regulatory purposes. With CASAC concurrence, the Criteria Document concluded that properly conducted monitoring can provide an adequate index of population exposure to ambient air pollution in epidemiological studies (U.S. EPA 1996a; Chapter 7). Although personal monitoring may be practical for some occupational and epidemiological studies and has been employed in some past studies of air pollution, it is not realistic to require personal monitors in air pollution studies of daily mortality, which require urban scale population data over a period of years. In addition, the central monitor measurements are related to day-to-day variations in population exposures to ambient sources of PM, which is a more appropriate metric for setting a national ambient air quality standard. Most personal monitoring done to date does not permit separation of PM derived from outdoor air pollution from that derived from indoor and personal sources (e.g., smoking, cooking, cleaning). As discussed more fully in Appendix D, this kind of combined PM exposure information would actually be less relevant to the assessment of the effects of ambient air pollution than are outdoor measurements.

EPA acknowledged and responded to observed cross-sectional comparisons of ambient and personal PM data in Chapter 7 of the Criteria Document and the proposal notice. As described therein, the major premise underlying commenters' arguments on this issue is incorrect. The question is not whether central monitoring site measurements contain a signal reflecting actual exposures to total PM from both outdoor and indoor sources at the individual level; the relevant question is whether central monitoring site measurements contain a signal reflecting actual exposures to ambient PM for the subject population, including both ambient PM, while individuals are outdoors, and ambient PM that has infiltrated indoors, while individuals are indoors. The PM standards are intended to protect the public from exposure to ambient PM, not PM generated by indoor or personal sources. There is ample evidence, as discussed in Chapter 7 of the Criteria Document, that personal exposure to ambient PM, while outdoors and while in indoor micro-environments, does correlate on a day-to-day basis with concentrations measured at properly sited central monitors (U.S. EPA, 1996a, p. 1-10). EPA has, therefore, concluded that it is reasonable to assume that a reduction in ambient PM concentrations will reduce personal exposure to ambient PM, and that this will protect the public from adverse health outcomes associated with personal exposure to ambient PM.

(2) *Comment:* In direct contrast to the comment responded to above, some commenters have argued that the PM epidemiological results are confounded because the weather and other factors that cause daily variations in outdoor pollution will cause similar daily variations in exposures to indoor generated air pollution.

Response: EPA disagrees with this assertion. For this to be true, *outdoor ambient* pollution concentrations would have to be correlated with personal exposure to *indoor-generated* air pollution such as that from smoking, cleaning, and cooking. This argument is logically inconsistent with the other comments on the lack of any such correlation with personal exposure, and these commenters have offered no scientific evidence to support their claim. In response, EPA has performed conceptual as well as quantitative analyses of the relevant information from the PTEAM exposure study and finds no evidence for such a correspondence in the actual data (see Appendix D). Indoor-generated PM appears to be independent of and uncorrelated with the ambient PM measurements in the backyards of each home in the PTEAM study. Moreover, the PM data also call into question the suggestion that individuals tend to engage in behavior that increases their exposure to indoor air pollution as outdoor air pollution levels increase.

As documented in Chapter 7 of the Criteria Document, time-series community studies observed the effects of varying levels of ambient air pollution; therefore, the effects of indoor-generated air pollution would be independent of and *in addition* to the effects found in these studies. Commenters apparently believe EPA is claiming such studies are detecting the effects of total PM personal exposure. This misunderstanding is evidenced, for example, by Wyzga and Lipfert's (1995) treatment of the difference between ambient monitors and actual personal exposures as "exposure errors"; Brown's comment for API (IV-D-2247) that "if (ambient) PM is causally related to mortality/morbidity, then it is

personal PM exposure that must be reduced to have an effect." On the contrary, it is personal exposure to *ambient* PM that must be reduced to address the risk identified in community air pollution studies. Any lack of significant correlation between outdoor PM concentrations and personal exposure to total PM from all sources is irrelevant, except to the extent it may decrease the power of time-series studies to detect the effects of ambient pollution.

(3) *Comment:* Commenters have also raised theoretically based concerns on a related issue, namely errors in the measurement of the concentrations of air pollutants.

Response: See section II.B of the preamble to the final rule, and the detailed discussion in Appendix D.

b. Mechanistic and toxicological evidence

Comment: A number of commenters pointed out the lack of demonstrated biological mechanisms derived from the controlled human, animal, and other laboratory studies of components of PM. Some have said that, absent clear mechanistic support for biological plausibility, EPA should not proceed to revise the standards, based solely on the available epidemiology. In some cases such commenters make reference to the evaluative criteria of Hill (1965).

Response: As noted in the preamble to the final rule, with respect to biological plausibility, Hill noted that "this is a feature I am convinced we cannot demand. What is biologically plausible depends upon the biological knowledge of the day" (Hill, 1965). This statement is clearly pertinent to the toxicological and mechanistic understanding of the effects of PM and associated air pollutants, especially at lower concentrations. It is also important to stress that while the mechanistic evidence published as of the time the Criteria Document closed does not provide quantitative support for the epidemiological results, neither can such limited evidence refute these findings (U.S. EPA, 1996a; p. 13-27 to 28). In fact, our understanding of biological mechanisms for PM pollution effects is not sufficient to explain the effects observed at much higher concentrations in air pollution episodes, for which causality is generally accepted. Moreover, the toxicological literature has only recently begun to examine animal models (or controlled human studies) that might reflect the sensitive populations in question (the elderly, individuals with chronic respiratory and cardiovascular disease) or that adequately reproduce all of the physico-chemical properties of particles in the ambient atmosphere. In short, the absence of evidence of a particular mechanism is hardly proof that there are no mechanisms that could explain the effects observed so consistently in the epidemiological studies.[7]

---

[7] The absence of biological mechanisms did not deter CASAC from recommending revisions to the PM standards in 1982, in 1986, and again in 1996. The length of time it has

As discussed in the Staff Paper, under ideal circumstances, animal toxicology and controlled human exposure studies can provide qualitative and quantitative support for environmental epidemiology. In the case of PM, however, the lack of published experimental human and laboratory animal studies involving relevant exposure levels and experimental subjects representative of sensitive subpopulations identified in the epidemiological studies presents problems in providing an integrated assessment (U.S. EPA, 1996a; p 13-2). Epidemiological studies describe relationships between regionally and temporally variable mixtures of particles and gases in community air pollution and mortality and morbidity in sensitive populations -- most notably the elderly and individuals with cardiopulmonary disease, which includes adults and children with asthma.

In contrast, experimental studies of PM effects in humans tend to use healthy young adult humans (or those with only mild disease) and examine mainly reversible physiologic and biochemical effects from exposure to laboratory-generated acidic aerosols, sulfates or nitrates. Similarly, experimental studies on laboratory animals have tended to use genetically homogenous healthy animals to examine a broader range of effects from individual components of the PM mix. In both controlled animal and human studies, the limited number of individuals exposed greatly limits the ability to detect effects at concentrations close to ambient levels. For example, at the very high PM levels observed during the worst London episode (i.e. >1000 $\mu g/m^3$), the increase in mortality rate was such that about 2 in 10,000 London residents expired in a given 24-hour period. Obviously, directly relevant controlled human studies of this effect are out of the question, suggesting the use of animal studies. If animals have similar susceptibility to the average humans in this population, however, experiments to detect such effects would be prohibitively expensive because of the large numbers that would have to be exposed to detect an effect. This means animal models that mimic the greater sensitivity to PM effects observed in susceptible human populations are important in furthering research, as are more sensitive indicators of processes that could lead to more serious effects. In general, however, extrapolation of quantitative and qualitative results from animal studies to human is encumbered by methodological difficulties from differences in dosimetry. The various species used in inhalation toxicological studies do not receive identical doses in comparable respiratory tract regions when exposed to identical aerosols. Consequently few laboratory experiments have used appropriate models of susceptibility to PM, which limits evaluation of possible mechanisms and potential quantitative effects comparisons.

However, at least qualitative support for some of the epidemiologic observations has been reported for specific components of the ambient particle mix in controlled clinical studies of humans as well as studies in animals. Although results are reported only for

---

taken to uncover mechanisms underlying the effects of cigarette smoking serves as an important caution to those commenters who suggest that regulations should not proceed absent mechanistic understanding.

levels generally higher than those observed in the ambient air, for such studies, the biological responses occurring in the respiratory tract following PM inhalation encompass a range of effects including: respiratory symptoms such as wheeze and coughing, changes in pulmonary function, altered mucociliary clearance, inflammation, changes in lung morphology and tumor formation (U.S. EPA, p. 13-70, p. 11-1). Most of these changes are consistent with effects observed in the epidemiological, but at generally much lower concentrations.

For additional discussion of this issue, see Appendix A.

c. Issues related to causality and the consistency/coherence of the evidence

i) Comments on causality

Many commenters agreed with EPA's interpretation of the epidemiological data in the Criteria Document and Staff Paper as clearly indicating a need to revise the current PM standards. A number of commenters, however, citing accepted criteria used in evaluating epidemiological studies to assess the likelihood of causality (most notably those of Sir Austin Bradford Hill, 1965), disagreed with EPA's assessment of the likelihood that air pollution containing PM is causally linked to observed health effects.

In general, the commenters and their consultants applied these criteria to subsets of epidemiological studies evaluated in the Criteria Document. In addition to the general response to these comments contained in Section II.B of the preamble to the final regulation, EPA has prepared more detailed responses to the papers and reports that were submitted as appendices to the comments of some groups. The major points made in these submissions and EPA's detailed responses to these commenters are presented in Appendix A. The response below addresses one of the comments most frequently made regarding causality.

(1) *Comment:* Most of the above commenters placed heavy reliance on one of Hill's criteria for inferring causality, namely the strength of the association. More specifically, such commenters argued that, in the absence of a demonstrated biological mechanism, the relative risks of effects in the PM epidemiological studies are too low (less than values variously cited as 1.5 to 2.0) to reach any conclusions regarding causality or to form the basis for regulations. Many gave examples of spurious statistical correlations as evidence that not all associations are causal.

Response: While Hill appropriately emphasized the strength of the association (e.g., size of the relative risk) as important, he also pointed out that "We must not be too ready to dismiss a cause-and-effect hypothesis merely on the ground that the observed association appears to be slight. There are many occasions in medicine when this in truth is so" (Hill, 1965). EPA believes that the effects of air pollution containing PM is such a case. The relative risks at the lower concentrations observed in the more recent epidemiological studies still imply very substantial numbers of sensitive individuals

effected. Moreover, unlike the "textbook" examples of unlikely significant associations provided by some commenters (e.g., ice cream consumption correlated with heat stroke), the abundant epidemiological literature on combustion particles documents numerous occasions in which single short-term episodes of high air pollution produced unequivocally elevated relative risks. For the week of the well documented 1952 London air pollution episode, for example, the relative risk of mortality for all causes was 2.6, while the relative risk for bronchitis mortality was as high as 9.3 (Ministry of Health, 1954). Hospital admissions also increased by more than a factor of two. British epidemiologists in the 1950s concluded that increased mortality was likely when PM (as mass calibrated British Smoke <4.5 μm in aerodynamic diameter) exceeded 500 $\mu g/m^3$ (Martin and Bradley, 1960). This is only about a factor of 3 higher than that allowed by the current PM standard. Unlike the "textbook" and other unlikely statistical associations noted by some commenters, where the *only* evidence is for low relative risk, clear and convincing links between high-level PM concentrations and mortality and morbidity buttress the findings of similar associations at much lower PM concentrations as suggested in the more recent epidemiological literature.

These commenters also appear to ignore several epidemiological studies conducted at low PM concentrations in U.S. and European cities, including both short- and long-term exposures to PM air pollution, that find statistically significant relative risks of respiratory symptom categories in children in the range of 1.5 to 5 (see table below). Concentrations in these studies extend from moderately above to well below those permitted by the current PM_{10} standards. While most of the recent epidemiological studies of mortality and hospital admissions report comparatively small relative risks, the findings of relative risks well in excess of the 1.5 to 2 criterion cited by commenters (e.g., Gamble and Lewis, 1996) for earlier studies of high PM episodes, as well as the relative risks of 1.5 to 5 reported in more recent studies of less serious, but still important effects categories, lend credibility to EPA's interpretation of the results.

RECENT U.S. PM STUDIES WITH RELATIVE RISKS >1.5 TO 2

Study	Relative Risk (95% Confidence Interval)	Concentration Difference / PM Indicator	Mean (max) PM level	Health Endpoint
Long-term Exposure Studies (1 or more years)				
Dockery et al., (1989)	5.39 (1.00-28.6) 3.26 (1.13-10.28) 2.93 (0.75-11.60)	50 μg/m^3 PM_{15}	(20.1-59) μg/m^3 *	cough bronchitis lower respiratory symptoms
Ware et al., (1986)	2.80 (1.17-7.03) 2.75 (1.92-3.94) 2.14 (1.06-4.31)	100 μg/m^3 TSP	(39.9-114) μg/m^3*	bronchitis cough lower respiratory symptoms
Dockery et al., (1996)	1.65 (1.12-2.42)	7 μg/m^3 sulfate	4.7 (7.4) μg/m^3	bronchitis
Short-term Respiratory Morbidity Studies				
Pope et al. (1991)[1]	7.03 (1.55-31.98) 5.24 (1.54-17.86) 2.51 (1.74-3.63) 1.94 (1.13-3.33)	184 μg/m^3 PM_{10}	46 (195) μg/m^3	extra medication use regular medication use lower respiratory symptoms upper respiratory symptoms
Pope and Dockery (1992)[2]	2.03 (1.37-2.99) 1.93 (1.29-2.89) 1.68 (1.13-2.50)	100 μg/m^3 PM_{10}	76 (251) μg/m^3	upper respiratory symptoms cough lower respiratory symptoms
Neas et al. (1995)[3]	1.71 (1.16-2.50)	15 μg/m^3 $PM_{2.1}$	24.5 (88.1) μg/m^3	cough
Schwartz et al. (1994)	1.53 (1.20-1.95)	30 μg/m^3 PM_{10}	30 (117) μg/m^3 **	lower respiratory symptoms

* Annual mean range from highest to lowest concentration across cities
** Median concentration all cities combined

1. Results from school-based sample
2. Results from symptomatic sample
3. Weighted by proportion of hours spent outdoors during prior 12 hours

ii) Consistency of the Epidemiological Studies

Many comments received explicitly agreed with EPA's finding that a large body of compelling evidence demonstrates that exposure to particulate matter pollution is associated with premature death, aggravation of heart and lung diseases, increased respiratory illness and reduced lung function. They agreed with EPA that these studies present a consistent and coherent relationship between exposure to PM and both mortality and various measures of morbidity. By contrast another substantial body of commenters asserted that the epidemiological evidence on PM is not as consistent and coherent as EPA has claimed, and, in particular, charged

that EPA ignored or downplayed a number of studies that the commenters argue contradict the evidence the Agency cited as supporting the consistency and coherence of PM effects.

The studies, all of which commenters contend do a better job of addressing one or more key issues, such as confounding pollutants, weather, exposure misclassification, and model specification, than earlier studies, include (1) several that were available during preparation of the Criteria Document, and (2) a number that appeared after the Criteria Document and Staff Paper were completed. Because the status of the later studies differs from that of the earlier ones, for purposes of decisions under Section 109, the two categories are discussed separately in the preamble to the final rule and in the responses presented below. An overview discussion of EPA responses to these latter comments is contained in the preamble to the final regulation, and more detailed responses to the major themes with respect to studies available for the criteria review are contained in Appendix A and Appendix B, and, with respect to more recent studies, in Appendix C. The remainder of the discussion presented here focuses on particular comments with respect to groupings of studies that EPA examined in the criteria and standards review.

(1) *Comment:* Some commenters, who argued that sufficient evidence does not now exist to establish standards for $PM_{2.5}$, raised specific questions about the consistency of the findings apparent in $PM_{2.5}$ and related epidemiological studies considered by EPA. For example, API (IV-D-2247) observed that "five of six $PM_{2.5}$ morbidity studies did not find a statistically significant association between $PM_{2.5}$ and adverse health effects." In addition, commenters asserted that EPA used inappropriate and uncertain ratios of $PM_{2.5}$ to PM_{10} in setting the standards.

Response: As stated in the Staff Paper, and consistent with CASAC recommendations for $PM_{2.5}$ standards, the criteria and staff review have found that sufficient scientific evidence exists to warrant establishment of separate standards for fine particles. While the decision on the need to revise the standards is based on the full range of the over 80 key epidemiological studies cited in Chapter 12 of the Criteria Document, over 30 such studies included some measure of fine particles (including fine mass, BS, COH, acids, or sulfates) and most of these studies had significant findings for the fine particle indicator. As noted in Section II.F of the preamble to the final rule, nine of these studies were determined by EPA to be key studies for quantitative assessment of health effects related to directly measured fine particle concentrations, and these studies are listed in Tables V-12 and V-13 of the Staff Paper. As described in detail in the Criteria Document and Staff Paper, the nine key fine particle studies include gravimetric measures of fine particles ($PM_{2.5}$ or $PM_{2.1}$), and in some cases, fine particle components (such as sulfates or acids). The studies include one short-term exposure mortality study (Schwartz et al., 1996), one short-term exposure hospitalization study (Thurston, 1992, 1994), three short-term exposure studies of changes in symptoms or lung function (Schwartz et al., 1994; Ostro et al., 1991; Neas et al., 1995), two long-term exposure mortality studies (Dockery et al., 1993; Pope et al., 1995), and two long-term exposure morbidity studies (Dockery et al., 1996; Raizenne et al., 1996). In each of the nine studies, statistically significant relationships were found between fine particles and/or fine particle components and

adverse health effects.

A number of commenters made statements similar to the comments made by Swidler & Berlin (IV-D-2519) and Lehigh Portland Cement Company (IV-D-3488) that refer to a specific number of epidemiological studies. Although shorter lists can be derived from tables or Figures (e.g. Figure 1 in the proposal) that EPA has used to show studies using single PM indicators (e.g PM_{10} or $PM_{2.5}$), in examining the weight of the evidence regarding the need to revise the PM standards, EPA has generally relied on the full set of epidemiological studies in the CD. Contrary to some commenters' statements, there are more than just two studies of health effects that measured fine particle concentrations.

EPA recognizes that, due to the large number of studies reviewed and the multiple tables and figures included in the Staff Paper, there has been some confusion regarding the specific list of studies upon which EPA relied in deciding to propose revisions to the PM standards. To assist in identifying the key epidemiological studies cited in summary tables in the Criteria Document, a chart of the 87 key studies has been prepared and is included in Appendix B. Nevertheless, the decision to propose new PM standards was not based on a simple count of studies: rather, it involved a detailed consideration of the whole body of scientific literature so that a decision could be made on the weight of the evidence for health effects related to $PM_{2.5}$. While the full body of studies was used in reaching the decision to revise the standards, in deciding upon the specific levels for the $PM_{2.5}$ standards, EPA placed greatest weight on the nine fine particle studies noted above.

API specifically refers to the $PM_{2.5}$ morbidity studies, stating that five of six studies do not show significant effects for $PM_{2.5}$. EPA believes the written record of these studies disproves this assertion; all of the morbidity studies that measured gravimetric fine mass have significant findings for fine particles or fine particle components. Although the point is not specifically discussed by the commenter, EPA believes that, for the purpose of assessing the health effects evidence, it is unreasonable to assert any significant distinctions between gravimetric measurements of $PM_{2.5}$ and $PM_{2.1}$. Both are clearly gravimetric indicators of fine mass. Moreover, where studies found statistically significant associations with $PM_{2.5}$ components (e.g., sulfates and/or acids, in Thurston et al., 1994; Dockery et al., 1996), EPA believes it is appropriate to use the corresponding $PM_{2.5}$ or $PM_{2.1}$ values from the study as an index of fine particle levels of concern. As noted in other responses and in the Criteria Document and Staff Paper, it is clearly difficult to separate the effects of subcomponents from the overall fine mass indicator. This approach is consistent with the advice of CASAC panelists who recommended the citation of fine PM component studies in the key quantitative study summary tables in the Criteria Document and Staff Paper. One of the clear outcomes of implementing $PM_{2.5}$ standards is the reduction of such major components of $PM_{2.5}$ as sulfates and acids. EPA also notes that the approach it has followed to selecting the level of the fine particle standards involves no conversions from the original measurements of fine particles used in these studies

With respect to the 6 morbidity studies, significant associations were found in the Six Cities study (Schwartz et al., 1994), as noted by the commenter, for both cough and lower respiratory symptoms with $PM_{2.5}$. Neas et al. (1995) report significant increases in incidence of evening cough episodes with $PM_{2.1}$, sulfate particles and particle-strong acidity. In addition, these investigators found significant decreases in childrens' peak expiratory flow rate with both sulfate particles and particle-strong acidity, which are both components of fine particulate matter, although no association with $PM_{2.1}$. In a study by Ostro et al. (1991) in Denver, significant associations were found for cough and shortness of breath with exposure to acid aerosols, and there was a significant association between exposure to sulfates and shortness of breath. The authors also report a positive, but not significant, association between fine particles and increased asthma rating ($p<0.10$) (Ostro et al., 1991). Thurston et al. (1994) also found respiratory hospital admissions to be increased significantly with increases in aerosol acidity, $PM_{2.5}$ and sulfate concentrations, when examined individually, but only the acid component of fine mass remained significant in multiple pollutant analyses; the findings of this study are discussed in greater detail in a separate response. In the 24-city study, $PM_{2.1}$, sulfates and aerosol acidity were associated with reduced lung function (Raizenne et al., 1996) and sulfates and aerosol acidity were significantly associated with increased bronchitis (Dockery et al., 1996).

In addition to the six key morbidity studies, one short-term (Schwartz et al., 1996) and two long-term exposure (Dockery et al., 1993; Pope et al., 1995) mortality studies included direct gravimetric measurement of fine particles. Again, each of the three found significant associations between exposure to fine particles and mortality. The short-term exposure mortality study (Schwartz et al., 1996) found increases in risk of mortality with $PM_{2.5}$ exposure, and the relative risks in this study ranged from 1.020 to 1.056 in the six cities under study. On an individual city basis, three of the four cities with annual averages of $PM_{2.5}$ above 15 μg/m3 found statistically significant associations between mortality and $PM_{2.5}$. The fourth city, Steubenville, was nearly significant for $PM_{2.5}$ and significant for $PM_{10.}$ Given that the relative risk is in line with those from the other more populous cities, the most likely explanation for lack of full significance is the fact that Steubenville had a much smaller population (by a factor 6 to over 10) than the other three. The long-term exposure mortality studies (Dockery et al., 1993; Pope et al., 1995) found relative risks ranging from 1.10 to 1.46 for mortality with increased exposure to $PM_{2.5}$ or sulfates. In the long term six city comparison of similar sized cohorts, Steubenville had a significantly elevated risk of mortality. Although EPA views the quantitative results of these long-term studies as more uncertain, they provide substantial support and insights into the potential nature of fine particle effects.

These nine studies were listed in the Staff Paper as key studies of the relationship between health effects and fine particle concentrations. As described above, each study reports significant findings for fine particles and/or fine particle components and increased mortality or morbidity. In its consideration of the need for additional PM standards, EPA weighed the findings of these key studies along with the results of other

studies, including those using particle measures such as BS or COH. From this weight-of-evidence approach to reviewing the scientific literature, EPA has concluded that there is consistency and coherence even in the more limited number of studies that examined indicators of fine particles (EPA, 1996b; p. V-76).

In selecting the levels of the $PM_{2.5}$ standards, EPA relied most heavily on these studies, which actually measured fine particles, and not on estimates derived from the use of uncertain ratios. Because of the qualitative and quantiative consistency observed in the PM studies in general, however, EPA notes that if such ratios were to be used as a basis for selecting the levels for $PM_{2.5}$ standards using studies that measured PM_{10}, it is likely that the resultant standard levels would be in the same ranges as were consdered using the $PM_{2.5}$ studies.

(2) *Comment:* UARG (IV-D-2250) observed that, of the 13 study locations highlighted in the review of short-term mortality studies, 8 have relative risks for which the confidence interval includes the null result.

Response: EPA does not agree with the commenter's implication that most short-term mortality studies cited in the Criteria Document have nonsignificant results. As indicated in Appendix B of this document, 90% of the key 87 studies upon which EPA relied for this decision had statistically significant results for some or all of the health endpoints used in the study. In this specific comment, reference is made to the 13 studies of short-term exposures to PM_{10} and mortality that are listed in Table V-3 of the Staff Paper. Even here, 9 of these 13 studies had statistically significant findings reported by the investigators, with the remaining showing positive results that are at or near significance.

The table in question includes all mortality studies for which a quantitative comparison could be made with respect to relative risk per unit daily increase in PM_{10}, in this case 50 $\mu g/m^3$. The results shown in the table are the results of calculations made by EPA for the Criteria Document, in which the original results are transformed into this common metric, which in this case is rounded to two places after the decimal. As noted by the commenter, this results in five of the associations having confidence intervals that include 1.0, which is itself on the borderline of significance at the 95% level. This, however, is largely the result of rounding of the calculated values for the table; the original results for all five studies were reported to be statistically significant by the original investigators. EPA calculated a lower bound confidence limit of 1.005 for 3 studies and 1.001 for the fourth, which are all above 1.0, but these were rounded to 1.0 in the table. In the fifth study (Kinney et al. 1995), the lower confidence limit of relative risk for PM_{10} alone is reported as 1.00, which as noted above, is at the threshold for significance.

Of the three remaining associations with confidence intervals that encompass 1.0, the original researchers report positive but not statistically significant results. One of the six cities, Portage, had a positive and nearly significant association (C.I.= 0.98 to 1.09),

while a second, Topeka, was clearly not significant (C.I. = 0.90 to 1.05). The remaining positive but not significant result for Kingston, TN reported in Dockery et al. (1992) was for a study of about one year duration. As noted in Chapter 12 of the Criteria Document, such a short duration limits the statistical power in a relatively small study population. Support for this suggestion is provided by a follow-up study of mortality (Schwartz et al., 1996), which used multiple years of data from the same site and an expanded population and found a statistically significant association.

Although some studies report findings that are not statistically significant or of border line significance, EPA reiterates the observation that there is remarkable consistency in the positive associations found from one location to another. EPA believes that in focusing on the statistical significance threshold and ignoring the clear patterns of associations revealed even in Table V-3, commenters are ignoring the guidance of recognized epidemiological experts (Greenland, 1991). When the scientific literature on PM-related health effects is considered as a whole, it is clear that ambient PM is clearly associated with serious health effects.

(3) *Comment:* ATA (IV-D-2245) lists a "core database of 36 epidemiological studies" that were used by EPA, and states that over one-third of these studies showed no statistical association between PM and adverse health effects.

Response: The characterization of the studies listed in Exhibit 10 of ATA's comments as EPA's "core database" is erroneous and misleading, and the specific listings are both incomplete and contain some inaccuracies. The studies listed are all drawn from Table 12-2 in the Criteria Document, which lists epidemiological studies of short-term exposure to PM and mortality. Thus, the commenter has selectively chosen to include only studies that use mortality as an end-point, and only short-term exposure studies. The Criteria Document, the Staff Paper, and the proposal notice also clearly took into consideration results of studies that used other measures of health, including admissions to the hospital or emergency room, and changes in lung function or respiratory symptoms. An examination of Chapter 12 of the Criteria Document includes additional Tables that address these effects categories, which include Tables 12-8, 12-9, 12-10, 12-11, 12-12, 12-13, 12-16, 12-21, and 12-22.

Taken together, a total of 87 separate epidemiological studies are listed. A separate tabulation of the above studies is included in Appendix B to this document. As shown in this summary table, the vast majority of the studies had results indicating deleterious effects of PM on health; 68 reported statistically significant associations and 11 studies had "mixed" findings.[8] Eight of the listed studies found no significant associations

[8]As discussed in Appendix B, the "mixed" category includes studies where results with respect to PM effects are less clear, including, for example, multi-pollutant studies in which the authors noted difficulties in separating the effects of PM from other pollutants, even if PM is

between PM concentration and health, with none of the studies finding consistent statistically significant negative results (indicating that PM is protective of health). In other words, there were no studies reporting significant reductions in mortality or hospitalization or significant improvements in lung function with increases in PM concentration.

EPA also notes several inconsistencies between the ATA partial listing of studies and the corresponding table in the Criteria Document. Table 12-2 actually lists 38 studies, not 36. ATA omitted 3 of those studies (Ito et al., 1993; Katsouyanni et al., 1996; Saldiva et al,. 1995) and apparently counted Styer et al. (1995) as two separate studies by independently considering the results for Cook and Salt Lake Counties. If it is appropriate to consider the results of each city examined by Styer et al. (1995) as separate studies, then the results of Ito et al. (1995) for Los Angeles and Chicago should be treated similarly, as should the reports from the Six Cities study (Dockery et al., 1992; Schwartz et al., 1996). It is also not clear why three of the studies from the Criteria Document were excluded from the list in ATA's Exhibit 10. Of the three studies not included in ATA's list, two report significant associations with PM (Ito et al., 1993; Saldiva et al., 1995) while one has mixed findings (Katsouyanni et al., 1990b). Among the remaining studies, there is general agreement between ATA and EPA regarding findings of statistical significance. In two instances (Samet et al., 1996b; Wyzga and Lipfert, 1995b) the ATA designates the studies as "yes" for statistical significance, but they are considered to have mixed results by EPA. The reverse is true for two studies that are listed as statistically significant by EPA but designated as "no" by ATA (Touloumi et al., 1994; Ozkaynak et al., 1994). Of the remaining studies, six were found to have mixed results by EPA that were categorized as non-significant by ATA (Kinney and Ozkaynak, 1991; Moolgavkar et al., 1995a; Moolgavkar et al., 1995b; Li and Roth, 1995; Samet et al., 1996a; Xu et al., 1994). With respect to a summary for the subset of mortality studies listed by ATA, EPA concludes that 70% were positive and significant (similar to the 2/3 estimated by ATA), while 21% were mixed, and only 9% found no significant results.

With respect to the full set of 87 studies, 90% of those EPA relied upon in proposing new PM standards find at least some statistically significant associations, and 78% of the studies can be considered "fully significant" studies in that the PM associations remain significant when alternative analyses or multipollutant models are used. Only 8 of the 87 studies (9%) used as a basis for the health-based standard did not find statistically significant associations between PM and adverse health effects.

As discussed more fully in Appendix B and the response to the next comment, EPA examined the mixed and "negative" studies in an attempt to identify factors that could result in positive or mixed associations despite a lack of statistical significance. In many

---

significant by itself; e.g., Samet et al. (1996a,b); Moolgavkar et al., (1995b).

cases sample size (number of days or number of subjects) greatly limited statistical power (Figure 12-17; Table 12-25 ; U.S. EPA, 1996a). In the case of multiple pollutant studies, the criteria review notes the difficulty in isolating effects of PM from other pollutants in individual studies. However, EPA cannot accept the suggestion that such multi-pollutant studies are in any way "negative" with respect to EPA's conclusions that PM, alone or in combination with other pollutants, is associated with adverse effects at levels below those allowed by the current standards. This conclusion is based not only on the consistency of PM effects across areas with widely varying concentrations of potentially confounding copollutants, but also on the extended analyses of the Philadelphia studies in the Criteria Document and Staff Paper.

Based on its assessment of the full epidemiological literature, EPA is confident that the decision on the need to revise the PM standards is soundly based on a consistent set of epidemiological studies showing a multiplicity of PM effects in sensitive populations at levels permitted under the current NAAQS.

(4) *Comment:* Ford Motor Company (IV-D-5323) asserts that EPA "has neglected to seriously explore why many studies (e.g., about 25% for the morbidity studies) have failed to show a positive, statistically significant association."

Response: EPA notes that the commenter's observation indicates that approximately 75% of the studies on which the PM standards are based show positive, statistically significant associations. Of the remaining studies, EPA reiterates the observations noted in the preceding response -- that an additional 15% had near-significant or mixed positive results, while none of the studies had consistent, statistically significant "negative" associations, and only about 10% failed to find any significant or near significant "positive" associations. As noted in section II. B of the preamble and in Appendix B to this document, it is important to note that the somewhat artificial designations of "negative" and "positive" findings actually suggest findings of, respectively, protective and deleterious effects. When including studies that show mixed positive results, EPA has determined that 90% of the studies provide evidence for PM-related health effects.

EPA does not agree with the commenter's assertion that EPA has neglected to explore or explain reasons why studies do not find statistically significant associations. Extensive discussions of issues that may affect epidemiological studies were included in Section 12-2 (Methodological considerations, pp. 12-9 to 12-27) and section 12-6 (Discussion, pp. 12-255 to 12-363) of the Criteria Document, along with a separate subsection on methodological issues in the discussion of long-term exposure mortality studies (pp. 12-139 to 12-147). The ability to detect an association between an exposure and an effect, if the relationship truly exists, is often referred to as statistical power. Some of the factors that can affect the statistical power of a study, and that were discussed in the Criteria Document, are sample size, the effects of simultaneous exposure to several pollutants, and model specification.

The one factor most commonly addressed in increasing the statistical power of a study is sample size. As described in the previous response, EPA evaluated the relationship between the t-ratio and sample size (in days) for a number of time series studies of mortality, and found close correspondence (Figure 12-17, EPA, 1996b). The t-ratio was found to increase with increasing sample size; using these data, EPA determined that the minimum sample size necessary (at 80% power) to detect a significant association between PM and mortality in a time-series study is 800 days (two-tailed test at the 0.05 level). It is notable that a number of time-series studies of mortality and PM exposure have had smaller sample sizes than that calculated by EPA to be the minimum needed to detect significant associations. For example, the studies in Kingston and St. Louis (Dockery et al., 1992) and Los Angeles (Kinney et al., 1995) had fewer than 400 days of observations, and both studies found positive associations that were either nonsignificant or at borderline significance.

Similarly, the analysis of data in subsets can result in sufficient loss of power so that a nonsignificant result cannot be interpreted as indicating the lack of an association. A number of researchers analyzed associations for each season of the year, thus reducing the sample size for each separate analysis. This form of subset analysis may be entirely appropriate in areas where PM concentrations vary by season, but the potential reduction in statistical power must be considered in interpreting the results of the analysis. One example of a study using subset analysis was conducted using data from Cook and Salt Lake Counties (Styer et al., 1995). The data set included six years of daily records, or over 2000 days of observations, and the researchers conducted subset analyses by both season and month. In the month-by-month analysis, PM_{10} was selected as an important explanatory variable during May and September for Cook County data, and during June and July using Salt Lake County data. EPA is not aware of a basis for presuming that an association between PM and mortality would vary by month of the year (other than seasonal variations that are accounted for in a season-by-season analysis); such extensive subdivisions of a data set can result in greatly reduced power and may produce essentially meaningless results. In addition, some researchers have created pollutant variables by dichotomizing the data into “high” and “low” categories, or dividing into categories such as quartiles or quintiles. Again, this may be entirely appropriate for the analysis of a particular data set, and EPA is not being critical of all studies that create categorical pollutant variables, but it must be recognized that some of the data may essentially be lost when data are grouped in this manner.

The specification of variables in models is another factor that can influence the results of an analysis. Overspecification of a model can cloud the results so that true associations are buried in nonsignificant associations with extraneous variables. This issue is described in greater detail in Appendices A, B, and C.

As described previously, about 13% of the 87 health effects studies used as a basis for EPA’s decision were found to have “mixed” results; positive associations were found that were statistically significant in some models but not significant in others. For the most

part, the change that resulted in PM associations losing statistical significance was the addition of other pollutants to the models. A number of studies have found the effects of PM to be difficult to separate from associations with a co-pollutant; the particular co-pollutant of interest varies in different locations and with different health endpoints. The many analyses of mortality data from Philadelphia have found the effect of TSP difficult to distinguish from the effects of other pollutants, especially SO_2 (Samet et al., 1995), while hospitalization studies in Canada have found effects of fine particles that become nonsignificant when modeled with ozone (Thurston et al., 1994). The problem of addressing confounding or collinear pollutants was discussed at length in the Criteria Document, and EPA believes that this issue can only be addressed by considering studies from a variety of locations, as stated in Samet et al. (1996a): "Insights into the effects of individual criteria pollutants can be best gained by assessing effects across locations having different pollutant mixtures and not from the results of regression models of data from single locations." Indeed, associations from different study locations presented in several tables in the Staff Paper indicate the remarkable consistency of effects that has been seen, even where the associations are not statistically significant. Significant associations have been found in areas with low levels of SO_2 such as Spokane (Schwartz, 1996) and Utah Valley (Pope et al., 1992) and in locations in the U.S. and other nations with varying climates.

What is most striking about the scientific evidence on PM and health effects is the consistency that is seen between studies. The great majority of PM-related health studies report significant findings for deleterious health effects, and the magnitudes of the associations are consistent from one study to the next, even in those studies where the association is not statistically significant. In evaluating potential reasons for not reporting a significant association, it can be seen that many of the studies with nonsignificant or "mixed" results may suffer from low statistical power, or from difficulty in distinguishing health effects from concurrent exposures to multiple pollutants. EPA not only disagrees with the commenter's assertion that the Agency has failed to assess possible reasons for findings of nonsignificance, but also believes that some of the factors identified in the review of this issue may account for the reported nonsignificant results from some studies.

(5) Other comments on statistical significance and consistency noted that the pooled relative risk for mortality associated with short-term PM_{10} exposures based on studies conducted in 10 locations is not statistically significant since the lower bound of the credible interval is less than 1.0 in Table VI-2 of the PM Staff Paper.

Response: EPA reexamined the pooled relative risk credible intervals reported in Table VI-2 of the Staff Paper and discovered that the intervals had not been corrected to reflect the methodology recommended by CASAC and actually used in the final Abt Associates 1996 risk assessment report (Abt Associates, 1996b). The correct 95% credible interval for the pooled relative risk based on these studies is (1.01 - 1.07). EPA has placed a memo in the docket (Richmond, 1997) that includes a corrected version of Table VI-2

reflecting the credible intervals reported in the final Abt Associates 1996 risk assessment report.

iii) Coherence

(1) *Comment:* Asthma incidence in the U.S. has been increasing in recent years despite the fact that PM levels have been decreasing. This would be inconsistent if PM were a cause of asthma.

Response: It is not believed, based on current evidence, that exposure to ambient PM is a major cause of asthma. The etiology of asthma is currently not well known, and numerous factors, such as family history of respiratory disease, have been shown to be associated with an increased risk of developing asthma. There is strong and convincing evidence, however, that exposure to air pollutants is associated with exacerbation of asthma. Therefore, EPA has considerable interest in asthma as a public health and as an environmental health issue.

As summarized in the CD and Staff Paper, increases in PM have been associated with increased hospitalization for asthma, worsening of symptoms, decrements in lung function and increased medication use; the available evidence does not demonstrate an association between PM exposure and asthma mortality. In the Staff Paper (p. V-35), EPA includes asthmatic individuals as one of the sensitive subpopulations that may be more susceptible to adverse health effects from exposure to ambient PM. A reduction in levels of PM or PM constituents is expected to reduce the number of asthma admissions to the hospital or the occurrence of asthma symptoms or medication use; EPA has not claimed that reducing PM pollution will necessarily reduce the incidence of new cases of asthma in a given year. It should be noted that, whether PM is a factor in causing an individual to develop asthma or not, the steadily increasing incidence of asthma means that a substantial portion of the population can be considered to make up the sensitive subpopulation. EPA believes that the scientific literature is supportive of a role for PM in exacerbation of asthma, and that this evidence along with evidence of other PM-related health effects provides ample reason to revise the standards for PM.

(2) *Comment:* The American Council on Science and Health (IV-D-2173) criticizes the use of "ecological" studies, and argues that there is insufficient evidence linking health effects to exposure to particulate matter. The commenters present a figure showing an overall decrease in PM_{10} concentration from 1987 to 1994, while there were small increases in mortality from COPD and asthma.

Response: Both Section II.B of the preamble to the final rule and Appendix A discuss the reasons EPA believes the available scientific evidence strongly supports the decision to revise the PM standards. In this case, the commenter ignores the strengths of time series studies of mortality or morbidity associations with PM exposure. These time series studies generally assess day-to-day changes in health with day-to-day changes in

pollutant concentration, while accounting for such factors as seasonal or other trends and confounding influences, if necessary. They are certainly not of the oversimplified study design demonstrated in the commenter's comparison of annual changes in PM_{10} concentration with annual mortality rates. In fact, epidemiologists have recognized that PM concentrations have decreased, and studies with multiple years of data have incorporated a long-term trend variable to account for this change. In addition, there has been a trend of decreasing rate in overall and cardiovascular mortality. In the most recent HEI analyses of the mortality-PM relationship in Philadelphia, these trends are depicted graphically, and the analysts describe several methods that were tested to control for such long-term trends (Samet et al., 1996a,b). After controlling for long-term trends, confounders and other factors, the authors found a significant association between exposure to air pollutants and mortality. Using the approach advanced by the commenter, EPA could claim that the observed trends in mortality and pollution are clearly coherent with an effect of PM on overall and cardiovascular mortality, and the much smaller rates of COPD and asthma changes are dominated by some other uncontrolled factors. However, EPA believes that it is more appropriate to note that the commenter's argument and examination of coherence is without merit because it relies on excessive oversimplifications and exclusion of factors known to be associated with these long-term trends.

d. New studies and analyses

As discussed in section II.B of the preamble to the final rule, a number of epidemiological and related studies, characterized as so-called "negative" evidence ignored by EPA, that were published or otherwise made available only after completion of the Criteria Document and Staff Paper for PM. EPA agrees that it did not rely on these studies, based on its long-standing view that the Act requires NAAQS decisions to be based on studies and related information included in the pertinent air quality criteria and available for CASAC review. See section II.B of the preamble to the final rule.[9] Although the Administrator has not relied on the more recent studies in reaching her final decision, the Agency has conducted a provisional examination of these and other recent studies to assess their general consistency with the much larger body of literature evaluated in the criteria review; the examination is presented in Appendix C. This assessment, while much less inclusive and rigorous than a criteria review, finds no basis for commenters' assertion that full consideration of selected new studies in the final decision would materially change the Criteria Document and Staff Paper conclusions on the

[9]Contrary to the views of one commenter, this does not necessarily limit the Administrator to consideration of information discussed in the criteria document for a pollutant (National Stone Association, IV-D-2999). To the extent that a staff paper or supporting document made available for CASAC review includes either scientific or technical information of the kinds specified in section 108(a)(2) of the Act but not evaluated in the corresponding criteria document, the Administrator considers that information to be part of the air quality criteria for the pollutant in question.

consistency and coherence of the PM data, or on the need to revise the current standards.

e. Health risk assessments

i. General methodology issues

This section addresses comments primarily summarized in Section II.A.3.f.(1) of the Summary of Comments document concerning various aspects of the methodology used to analyze health risks associated with alternative $PM_{2.5}$ standards. A number of industry commenters argued that EPA's risk assessments are flawed and incomplete and present an overestimate of the risks associated with PM exposure. Reasons cited by these commenters included: (1) criticisms about use of epidemiological studies which these commenters argued are inadequate for the reasons summarized in Section II.A.4.a. of this document, (2) disagreement with the assumption of a linear no-threshold concentration-response relationship, (3) disagreement with the basic assumption that $PM_{2.5}$ provides an appropriate indicator that is likely to be related to health effects, (4) concerns about the relative toxicity of components of $PM_{2.5}$ and whether or not reductions in $PM_{2.5}$ would result in reduction in the components responsible for any effects, (5) disagreement with the choice of background PM_{10} and $PM_{2.5}$ levels used in calculating risk in excess of background, and (6) concerns about lack of consideration of personal and indoor exposures in epidemiological studies used in risk analyses.

EPA notes that it is not the function of the risk analyses to address questions concerning causality, mechanisms, and related issues. Judgments about such issues are described elsewhere and the risk analyses show the results if one uses these judgments. Indeed, the risk analyses summarized in the Staff Paper and proposal notice and described in more detail in technical support documents (Abt Associates, 1996 a,b; Abt Associates, 1997 a,b) acknowledge these issues and uncertainties and illustrate the potential influence of many of these uncertainties in sensitivity and integrated uncertainty analyses. As discussed in Section II.B.2 of the preamble to the final rule, EPA believes that, even recognizing the large uncertainties, the key qualitative insights derived from the risk assessment and summarized in Section II.A.3. of the preamble to the final rule remain appropriate. While not placing great weight on the specific numerical estimates, EPA believes that the risk analyses confirm the general conclusions drawn primarily from the epidemiological results themselves, that there is ample reason to be concerned that exposure to ambient PM at levels allowed under the current air quality standards presents a serious public health problem. Presented below are responses to various specific issues related to the health risk analyses, expanding upon the discussion contained in the preamble to the final rule.

(1) *Comment:* The assumption of causality is not appropriately addressed in the risk analyses.

Response: The issue of causality in the PM epidemiological evidence is addressed in Section II.B.2 of the preamble to the final rule and in Section II.A.4.c. of this document. EPA continues to believe that the assumption of causality of PM across a range of PM concentrations, either directly or as a useful index for the mixture of pollutants related to

the health effects in question, is appropriately addressed in the Staff Paper discussion of the risk analyses (U.S. EPA, 1996b; p. VI-1). Moreover, in some specifications, including one illustrated in the proposal (Figure 2.c), the risk assessment assumed a threshold for the range of causal associations. See also next response.

(2) *Comment:* EPA inappropriately assumed linear, no-threshold concentration-response relationships were appropriate for its risk analyses. While the risk assessment included sensitivity analyses examining alternative potential thresholds, these estimates were not factored into the Agency's conclusions.

Response: As recognized by at least some of the commenters, EPA's risk analyses examined the impact on risk estimates of alternative concentration-response relationships that included various "thresholds" both in a series of sensitivity analyses and in an integrated uncertainty analysis. Contrary to the assertion made by some commenters that EPA ignored this information, the Staff Paper includes extensive discussion of the results of both the sensitivity and integrated uncertainty analyses (see Staff Paper, pp.VI-35-VI-43 and VI-54-VI-58). The Staff Paper discussion of key observations from the risk analyses (Staff Paper, pp.VI-58-VI-60) includes presentation of both base case (no threshold) and integrated uncertainty (various thresholds assumed) analyses estimates. Key observation number 4 in the Staff Paper (p.VI-59) states that, "Based on the results of the sensitivity analyses of key uncertainties and the integrated uncertainty analyses, the single most important factor influencing the uncertainty associated with estimates of PM health risk is whether or not a cutpoint concentration exists below which PM health risks are not likely to occur." Moreover, EPA displayed the results of the assumption of a potential threshold in the illustration of the distribution of annual risks associated with PM in Figure 2c of the proposal, and summarized the risk assesment conclusion on threshold both there and again in the preamble to the final rule. EPA has fully considered the uncertainty about the potential existence of a threshold in its final decision, as discussed in the preamble to the final rule.

(3) *Comment*: Environ International Corp., on behalf of Kennecott Corp (IV-D-2213) stated that the assumption of a log-linear no-threshold concentration-response function is not scientifically justified.

Response: The log-linear no-threshold model is the most common model for concentration-response functions in the literature on PM and health effects. In estimating the relationship between PM concentrations and a given health endpoint, researchers usually investigate the model form and choose a form that fits the data well. If there were clear evidence of a threshold, one would expect this to be reflected in threshold models having been chosen over log-linear models by researchers. There is not, however, clear evidence of a threshold. So far, there are insufficient data to determine whether a threshold model fits the data any better than a non-threshold model. Nevertheless, threshold models were examined in both the integrated uncertainty analyses and in sensitivity analyses. Because of lack of information concerning the

likelihood of different possible thresholds, however, analyses of thresholds relied on professional judgment.

(4) *Comment:* Inadequate control of confounders likely to lead to substantial overstatement of PM health effects.

Response: The issue of confounders and whether or not their treatment in the epidemiological studies used by EPA was appropriate or likely to introduce a bias resulting in overestimation of health effects is addressed in the Staff Paper, proposal notice, and in Section II.B. of the preamble to the final rule. The issue is also discussed elsewhere in this document and its appendices.

(5) *Comment:* EPA's risk analyses are flawed because they do not recognize the lack of correlation between ambient PM concentrations and personal PM exposures and fail to evaluate how high indoor PM exposures may be confounding the results in certain epidemiological studies.

Response: Concerns related to ambient versus personal exposures in the context of interpreting the available epidemiological studies that relied on fixed-site monitors are addressed in Section II.B.2 of the preamble to the final rule and in Sections II.A.4.a. and II.A.4.c. and Appendix D of this document.

(6) *Comment:* Several commenters stated that EPA's risk analyses, as well as the standards, are based on an unfounded presumption that if $PM_{2.5}$ is controlled, the actual "culprit" related to any health effects will be controlled. Some commenters argue that, because it is difficult to separate the effects of PM from gaseous pollutants in some studies, the most that can be said is that air pollution may be related to mortality and other effects, with the role of PM unknown. They further assert that EPA has not shown whether or not any risk reduction or benefits will occur as the result of reducing fine particles, and that the risk assessment ignored this issue.

Response: The underlying issue regarding the inclusion of various components in the $PM_{2.5}$ indicator is addressed in responses to comments above on the use of $PM_{2.5}$ as an indicator. However, EPA disagrees that the overall criteria and standards review ignored the implications of multipollutant studies or that the full risk assessment that comprises these documents ignored this issue. While EPA believes it is more likely than not that $PM_{2.5}$ and its components are related to serious health effects at levels permitted by the current PM standards, the alternative view that PM could be acting as a surrogate for pollutant gases or PM components was discussed and the implications of alternative hypotheses were examined in a qualiative risk assessment in Chapter VII of the Staff Paper.

In this qualitative assessment of the potential effectiveness of fine particles as a surrogate, EPA staff considered the results of various analyses of air pollution and

mortality in Philadelphia (Moolgavkar et al., 1995; Wyzga and Lipfert, 1995; Samet et al., 1995; 1996a,b), as summarized in the following excerpt.

> The CD evaluation of these multiple investigations concludes that for this single city example, it appears most difficult to separate independent effects of PM (as TSP) and SO_2, concluding that the relationship between these pollutants and mortality may be inherently non-linear (CD, p 13-57). Several clearly hypothetical explanations have been advanced to explain these results. The following qualitative assessment of several speculative, but plausible hypotheses (in italics), outlines the potential implications of these alternatives for the effectiveness of fine particle control as a surrogate:
>
> - *The complex relationship is a statistical artifact and only one of the pollutants is causally related. If the pollutant is PM,* then fine particle control would clearly be beneficial. *If the pollutant is SO_2,* which occurs at moderate levels in Philadelphia, reductions in local and transported SO_2 precursor control prompted by a fine particle standard would reduce health risk.[10]
>
> - *The relationship is real and due to increased penetration of an SO_2 complex carried on carbonaceous or other non-acidic particles.* Then local controls of primary fine particle combustion sources would likely reduce risks, because reducing the aggregate particle surface area (by reducing fine mass) is more likely to reduce dose than SO_2 reductions.
>
> - *The relationship is due to the association between SO_2 and acidic sulfates, which are the active agent.* In this case, fine particle controls are clearly beneficial.
>
> - *The relationship is due to the combined interactions of SO_2 and particles in different regions of the respiratory tract.* Again, control of fine particles would be beneficial.
>
> The staff does not have to accept any one of these hypothesized explanations as more likely to conclude that control of fine particles as a class appears to be a reasonable approach to reducing health risks in this particular example of potential confounding. It is also useful to note that, because of their relatively low surface area and origin, such a conclusion would not be as applicable to control of coarse fraction particles.

---

[10] As noted in section V.E of the Staff Paper, the evidence across multiple areas shows that PM is consistently associated with mortality in areas with high and low SO_2, making the second explanation unlikely.

> Although the above examples of alternative consequences of the use of fine particles as a surrogate are limited to PM and SO_2 interactions, some of these outcomes would extend to PM interactions with other pollutants as well. Given the large surface area of aqueous droplet and/or dry fine particles, as well as the multiplicity of similar effects caused by common gaseous pollutants such as ozone and related photochemical products and precursors, and NO_2 in addition to SO_2, direct or indirect interactions among these pollutants would not be unexpected (Section V.F.; CD, p 13-9.). Because ozone precursors, including NO_2 and volatile organic compounds, are also secondary particle precursors, it is reasonable to expect that the control of fine particles could also prompt control of local and regional sources of some of these precursors as well as SO_2. On the other hand, beyond the possibility of effects modifications in the body, the potential for gas/particle interactions between PM and CO is limited. It is also less clear that fine particle control would prompt significant additional CO control, the major contributors of which, mobile sources, are already subject to significant national reduction requirements. The rationale for concluding that the existence of PM effects is unlikely to be due to confounding by other pollutants is discussed in Section V.E. [U.S. EPA, 1996b; p VII-12 to VII-13.]

Staff also explicitly considered the potential concerns raised about ultrafine particles, as discussed elsewhere in these responses.

The above examples also illustrate why, based on current information, it is reasonable to conclude that control of fine particles as a group is likely to reduce those components of air pollution that are responsible for the observed associations between air pollution and serious health effects. The qualitative assessment shows this is the case whether, as EPA concludes is likely to be the case, PM is contributing directly to effects - alone or in combination with other pollutants, or PM is acting as a surrogate for other pollutants that are precursors to PM. It also indicates that EPA did consider this important risk assessment issue in an explicit examination of alternative possibilities. EPA believed this qualitative approach was more appropriate, given the greater likelihood that PM is playing a role in the observed effects.

(7) *Comment*: The Sapphire Group for AAMA (IV-D-2243) noted that based on the information provided in the Staff Paper on exposure and mortality (the risk ratio), it is not possible to duplicate the final results of the Agency's risk assessment upon which the standard is based.

Response: In the original risk analysis, PM levels were taken only down to the lowest observed level in the study (which was 9 $\mu g/m^3$ in Pope et al., 1995)[11]; if the annual

[11]As noted in the preamble to the final rule, the original risk analysis used in this example erroneously assumed that the lowest level in this study was an annual mean. In fact, it is the

mean is used, the change in PM would be from the lowest observed level of 9 to the annual mean. The annual mean in Philadelphia is rounded to 17 $\mu g/m^3$ in the Staff Paper, so the change in PM in Philadelphia would be: 9 - 17 = -8. The attributable fraction (i.e., the percent of all deaths attributable to PM) is calculated as [exp(PM coefficient * change in PM) - 1] -- i.e., the relative risk minus 1. The PM coefficient in Pope et al. (1995) is 0.006408. The attributable fraction in Philadelphia would therefore be 5% (4.997%). The annual mean in L.A. is rounded to 30 $\mu g/m^3$ in the Staff Paper. The change in PM is therefore calculated as: 9 - 30 = -21. The attributable fraction in L.A. would be 12.59%. The discrepancies between these attributable fractions and those in the Staff Paper (4.6% for Philadelphia and 11.9% for L.A.) are due to using the rounded annual means here. The actual means are 16.5 $\mu g/m^3$ in Philadelphia and 28.7 $\mu g/m^3$ in Los Angeles.

If PM levels were taken down to background instead of only to the lowest observed level in the study, then the change in PM in Philadelphia would be (3.5 - 16.5) = -13, and the change in PM in L.A. would be (2.5 - 28.7) = -27.2. The attributable fractions associated with these changes in PM would be 7.99% (in Philadelphia) and 16% (in L.A.). These attributable fractions are significantly higher than those reported (4.6% and 11.9%, respectively) because they correspond to larger changes in PM -- taking the annual mean down to background instead of only to 9 $\mu g/m^3$. A reported risk is always specific to a particular change in PM levels. The relative risks (or attributable fractions) reported in the risk assessment for Philadelphia and Los Angeles are completely consistent with the relative risk of 1.17 reported by Pope et al., 1995, for a change of 25 $\mu g/m^3$. They are simply specific to different PM changes. The risk assessment uses the concentration-response function identified in the epidemiological study it relies upon (e.g., Pope et al., 1995). One should be able to duplicate the results of the risk assessment using the information provided in the supporting documents.

(8) *Comment:* EPA's risk estimates for mortality associated with long-term exposure for Philadelphia and Los Angeles which are based on the concentration-response relationship developed from Pope et al. (1995) are in error because the value cited in Pope et al. as the mean of the mean concentrations across the 50 cities was actually the mean of the median values across these cities.

Response: EPA acknowledged this error in April 1997 and has placed in the docket June 5 and June 6, 1997, letter reports from Abt Associates that provide corrected risk estimates for mortality associated with long-term exposure that are now based on the mean of the median $PM_{2.5}$ concentrations across the cities examined in the Pope et al. (1995) study. EPA advised the CASAC of the error and its approach for addressing the

---

annual median, which is equivalent to a somewhat higher mean (e.g., 11 $\mu g/m^3$). This has been corrected in the updated risk assessment and Staff Paper tables that have been placed in the docket (Richmond, 1997).

issue in correcting the risk assessment and the Staff Paper. EPA also has placed in the docket a set of corrected tables and figures (Richmond, 1997) from the Staff Paper that are now based on the appropriate PM value from the Pope et al. (1995) study.

(9) *Comment:* Several commenters stated that the two areas included in EPA's risk analyses are not representative of national exposures to PM. More specifically one commenter (Ford Motor Company, IV-D-5323) argued that Philadelphia was a poor choice of city for which to conduct risk analysis because 3 out of 4 studies (Moolgavkar et al., 1995; Samet et al., 1995; Li and Roth, 1995) found that PM was not signficantly associated with mortality in Philadelphia. Also, it is indefensible to assume that Southeast Los Angeles County can adequately represent the Western U.S. or that Philadelphia County adequately represents the East.

Response: As discussed in section II.B of the notice to the final regulation and Appendices A and B, EPA disagrees with the above assessment of the Philadelphia data, which in any case relied on TSP and not $PM_{2.5}$. EPA selected Philadelphia County and Southeastern Los Angeles County because EPA was able to obtain relatively recent and fairly complete PM_{10} and $PM_{2.5}$ air quality data for these two areas. EPA reviewed its criteria for selection of cities for the risk analyses and its choice of these two areas with the CASAC at its February 1996 meeting. EPA has never claimed that either area completely represents all cities in either the western or eastern regions of the U.S. Rather, these two areas were chosen, given the constraints of air quality data availability, to illustrate the risks under current air quality and upon attaining alternative standards in two areas with quite different air quality patterns.

ii. Selection of studies used in analysis

This section addresses comments primarily contained in Section II.A.3.f.(1)b) of the Summary of Comments document.

Comment: EPA's risk analyses should not have relied on the population-based epidemiological studies for a variety of reasons including: actual exposures are not well understood, biological mechanisms not clear, concentration-response relationships not adequately characterized.

Response: See response to comments in Section II.A.4.a. regarding criticisms of the epidemiological studies used in EPA's risk analyses. As described in the Staff Paper (pp.VI-11 to V-14) and in the technical support document (Abt Associates, 1996a; p.47), the choice of studies to be used in the risk analyses was based on the Criteria Document tables of studies that were judged adequate by the Criteria Document to provide estimated concentration-response relationships for a variety of health endpoints associated with elevated PM_{10} and/or $PM_{2.5}$ exposures. The selection of studies was presented to the CASAC at its February 1996 meeting and drafts of the risk assessment methodology and the risk assessment results in the Staff Paper were reviewed at the May

1996 CASAC meeting, and subsequent to that meeting, by individual CASAC members reviewing staff revisions to the risk assessment.

iii. Calculation of excess risk relative to background

This section addresses comments primarily summarized in Section II.A.3.f.(1)c) of the Summary of Comments document concerning the background levels for PM_{10} and $PM_{2.5}$ used in the health risk analyses.

Comment: The background levels used in the PM risk analyses are unreasonably low.

Response: See responses to comments in Section II.A.3.g. of this document, which address criticisms of EPA's estimation of background concentrations for PM_{10} and $PM_{2.5}$. With respect to the choice of background concentrations for purposes of risk analyses, the CASAC concurred with EPA judgments presented in the Staff Paper that estimates of the annual average background level rather than a daily average (e.g., the maximum 24-hr level) be used since risks are aggregated for each day throughout the year (see Staff Paper, pp. VI-3 to VI-7). The midpoint of the estimated ranges for background cited in the Criteria Document (p. 6-44) were used in the base case risk analysis. The risk analyses also included sensitivity analyses that examined the impact of using the lower and upper bound values of the Criteria Document ranges for background on the health risk estimates. Finally, the integrated uncertainty analyses section of the risk assessment included uncertainty about background, again assuming that background fell within the range cited in the Criteria Document.

iv. Characterization of uncertainties

Comment: EPA's risk estimates are misleadingly precise and understate the degree of uncertainty because they do not address adequately the following issues: (1) whether or not $PM_{2.5}$ is the causative agent, (2) the effects of differential measurement error, (3) whether or not a threshold exists, (4) inadequate control of confounders in the underlying epidemiology studies.

Response: EPA's extensive evaluation of these issues in the Criteria Document and elsewhere in the Staff Paper is referenced in response to other comments on these issues and uncertainties. The risk assessment discusses all of these sources of uncertainty qualitatively and, when possible, assesses them quantitatively as well. For some sources of uncertainty, however, information on which to base a quantitative assessment was judged to be insufficient. Uncertainty about whether the associations between PM and various health endpoints are causal relationships, for example, was addressed qualitatively, and it was concluded, with the support of CASAC, that "the weight of epidemiologic evidence indicates that ambient PM exposure has affected the public health of U.S. populations" (PM Criteria Document, p. 13-27). Information on which to base a reasonable quantitative assessment of this uncertainty, however, was considered

insufficient. The uncertainty about possible thresholds was examined quantitatively, both in integrated uncertainty analyses and in sensitivity analyses. Because of lack of information concerning the likelihood of different possible thresholds, however, quantitative analyses of this uncertainty had to rely on professional judgment. To the extent that there is an "errors in variables" problem, the risk assessment can only discuss it qualitatively, because there is insufficient information to try to incorporate the associated uncertainty in any quantitative uncertainty analysis.

v. Public health implications

(1) *Comment:* EPA indicated in the proposal notice that most of the risk results from low level PM exposures rather than peak exposures. However, this is in contrast to experiences in well known air pollution incidences (e.g., London, Donora, and the Meuse Valley).

Response: Figures 2a,b, and c included in the proposal notice show that for a typical urban area the "low- to mid-range concentrations (e.g., 10-50 $\mu g/m^3$) account for the largest amount of estimated mortality risk on an annualized basis." As explained in the proposal notice, "Even though higher 24-hour concentrations, including peaks above 70 $\mu g/m^3$, clearly contribute more mortality per day than low- to mid-range concentrations, the much larger number of days within the low- to mid-ranges results in this interval being associated with the largest proportion of total risk." There is nothing inconsistent between this observation and the very different situations which occurred in Donora, London, and the Meuse Valley where extremely high PM levels occurred over a number of days. While the historic London episodes were quantitatively different from those assumed in the risk assessment, the record over 14 London winters indicates a continuum of effects down to the lowest levels. It is therefore likely that the cumulative increase in mortality calculated for all the days in the whole 14-year period would not be dominated by the more limited number of episode days.

(2) *Comment:* EPA's risk assessment for Philadelphia and Los Angeles provides additional support for a more stringent 24-hour $PM_{2.5}$ standard because a 25 $\mu g/m^3$ standard would prevent an additional 1,200 premature deaths, 500 hospital admissions for respiratory and cardiac causes, and 14,000 fewer cases of lower respiratory symptoms in children, compared to a standard of 50 $\mu g/m^3$.

Response: See discussion in Section II.F. of the preamble to the final rule.

f. Altitude and temperature corrections

Comments received on this issue were divided. A number of commenters supported EPA's proposal to eliminate these corrections for PM. A few States opposed the change because the lack of adjustment for very cold temperature in areas near sea level could make the standard more stringent. Some commenters were concerned that the proposed change would relax

protection afforded for areas at high altitude. A few commenters expressed concern that "sojourners" who visit high altitude area would have higher ventilation rates and receive reduced protection as compared to local residents whose ventilation patterns were more adapted to these conditions. EPA's responses to these comments are presented in section VI.A of the preamble to the final regulations.

g. Characterization of background $PM_{2.5}$ concentrations

A number of commenters, who used differing definitions for "background," expressed concerns that EPA was establishing standards at levels that approached or was below background. These comments are further summarized and responded to in the responses to comments on levels above.

5. Comments on related programs

A variety of public comments were received in this area, including recommendations for establishing significant harm levels for $PM_{2.5}$ and revising those for PM_{10}, establishing an intervention program similar to the one EPA has proposed for SO_2, and comments on nondegradation programs and other voluntary programs. The comments are summarized in section II.A.2 of the Summary of Comments. In addition to the responses contained in section II.E of the preamble to the final rule, EPA provides the following responses to specific issues.

(1) *Comment:* Commenters recommended that EPA establish significant harm levels for $PM_{2.5}$ and revise the significant harm levels for PM_{10}. A State government (IV-D-2335) suggested that EPA consider an intervention level program similar to that considered for SO_2 to address shorter than 24-hour exposures. Other commenters recommended that EPA consider voluntary emissions reduction programs in conjunction with standards.

Response: EPA agrees that significant harm levels and episode criteria need to be established for $PM_{2.5}$. The western U.S. may have distinct needs in this regard. To address episodic events, EPA intends to establish a significant harm level for $PM_{2.5}$ and associated guidance so States can develop appropriate emergency episode plans. EPA also plans to re-examine its significant harm levels with respect to PM_{10}. During the time these programs are under development, EPA will continue to use the existing limits,and will encourage other voluntary programs and incentives discussed by commenters. The significant harm and episode criteria will be included in the forthcoming proposed revisions to 40 CFR parts 51 and 58. In the interim, the existing PM_{10} emergency episode plans should be triggered by events of the magnitude raised in public comment.

EPA notes that many areas use voluntary programs or incentive programs in conjunction with standards with good success, and EPA encourages their use. For example, several areas operate voluntary woodstove curtailment programs to prevent pollution episodes during winter inversions.

(2) *Comment:* A commenter (IV-D-2095) suggested the consideration of nondegradation program for PM_{10}.

Response: To ensure an effective transition between the current NAAQS and the revised NAAQS, EPA will retain the current standard in the interim period. See Section VII of the preamble to the final rule.

B. Secondary PM Standards

1. General comments on proposed secondary standards

a. Considerations for setting secondary standards identical to primary standards

This section addresses significant comments on EPA's decision to establish the secondary NAAQS for $PM_{2.5}$ equal to the primary NAAQS for the purpose of addressing welfare effects, including visibility impairment, soiling, and materials damage. Comments on the proposed sectiondary NAAQS are discussed in section III of the preamble to the final rule.

(1) *Comment:* Several commenters suggested that EPA should set separate $PM_{2.5}$ secondary standards at concentrations below those of the proposed primary standards to protect against a range of welfare effects associated with PM, including visibility impairment, acidic deposition, vegetation effects, materials damage, soiling, nuisance, and safety concerns. Some argued that while such standards may not address all adverse visibility impairment due to regional variations in levels and composition of natural background PM, such standards would address a substantial amount of impairment, particularly in the eastern U.S. Some commenters also felt that a specific rationale for excluding acidic deposition from the secondary standard discussion should have been provided by EPA in the proposal.

Response: See Section III of the preamble to the final PM NAAQS for an extensive discussion regarding EPA's rationale for not establishing separate secondary standards at concentrations below those of the primary standards. EPA acknowledges the effects of PM and its precursors on the public welfare in addition to visibility impairment, including acidic deposition, vegetation effects, materials damage, soiling, nuisance, and safety concerns. However, EPA has determined in its review of relevant studies and new information that the available data do not provide a sufficient basis for establishing a separate secondary standard for effects other than visibility. As explained in the preamble and below, setting separate secondary NAAQS for visibility at concentrations below those of the primary standards is problematic for other reasons and EPA is adopting an alternative strategy to protect visibility.

A discussion of acidic deposition was not included in the PM secondary NAAQS proposal for two principal reasons. First, as discussed below, EPA had recently

examined the appropriateness of establishing separate secondary NAAQS to protect against acidic deposition effects in the latest secondary NAAQS reviews for sulfur dioxide and nitrogen dioxide, the two pollutants which, along with the products of their chemical transformation in the atmosphere, are the principal contributors to acidic deposition. Second, the 1990 Amendments to the Clean Air Act established a separate program specifically to address acid deposition under title IV. CAA section 404 also required EPA to study the feasibility of an acid deposition standard.

In EPA's April 21, 1993 final decision that revisions to the secondary standard for sulfur dioxide were not appropriate (58 FR 21351), EPA took into account the significant reductions in SO_2 emissions, ambient SO_2 concentrations, and ultimately the deposition of sulfur that is expected to result from implementation of the title IV acid rain program. EPA also noted that it would be prudent to await the results of several studies and research programs, especially those designed to monitor progress resulting from the implementation of title IV, those assessing the comparative merits of secondary standards, and the section 404 acid deposition standard feasibility study (58 FR 21357).

In EPA's October 8, 1996 final decision that revisions to the primary and secondary standards for nitrogen dioxide were not appropriate (61 FR 52852), EPA concluded that the available scientific and technical evidence in the Criteria Document and Staff Paper did not provide an adequate basis for setting a separate secondary standard for nitrogen dioxide to address the effects associated with nitrogen deposition and acidification. EPA recognized the significant uncertainties associated with developing a consistent relationship between varying concentrations of NO_2 in the ambient air and atmospheric deposition of NOx and ultimately the appearance of nitrogen in surface waters. EPA also noted that in its review of the "Acid Deposition Standard Feasibility Study: Report to Congress," the Acid Deposition Effects Subcommittee of the Ecological Processes and Effects Committee of the EPA's Science Advisory Board concluded that there was not an adequate scientific basis for establishing an acidic deposition standard. The report found that it was not appropriate to establish an acid deposition standard because of 1) the variable sensitivities of streams and lakes within the same region to acid deposition, and 2) the uncertainty inherent in selecting an appropriate level for such a standard. The same difficulties would apply to setting uniform national secondary standards to protect against acidic deposition.

The "total loadings" approach to reducing acid deposition that was adopted by Congress in title IV of the Act avoids the scientific difficulties referred to above, and it addresses the problem in the manner and to the degree Congress determined to be appropriate. For all the above reasons, EPA believes it is both infeasible and inappropriate to establish at this time a separate secondary standard for PM to address acidic deposition.

(2) *Comment*: There is no justification for setting any secondary standards for PM since the primary standards and a regional haze program will sufficiently address visibility and other welfare effects.

Response: Secondary standards are to be established to address any known or anticipated adverse effects on the public welfare associated with a criteria pollutant. It is clear that coarse and fine particles can cause adverse effects on visibility. The PM Criteria Document shows the unequivocal scientific evidence for these significant impacts on visibility. (See Criteria Document, Chapters 8 and 9.) Thus, EPA has determined that it is appropriate to establish secondary standards for $PM_{2.5}$. See section III.A of the preamble to the final rule.

(3) *Comment*: Several commenters supported setting secondary standards for PM equal to the primary standards.

Response: EPA agrees that secondary standards for $PM_{2.5}$ should be set equal to the primary standards, but emphasizes that in order to appropriately address the regional differences in adverse effects of PM on visibility, it is also essential to establish an effective new regional haze program for the protection of visibility in mandatory class I Federal areas. As noted in section III.A.1 of the preamble, regional reductions of emissions that impair visibility should benefit the public welfare by improving visibility and reducing other welfare effects, both within and outside class I areas.

(4) *Comment*: Several commenters supported establishment of a regional haze program in conjunction with secondary standards equal to the primary standards, while other commenters stated that such a program is not needed.

Response: EPA believes a regional haze program under section 169A of the Act will be a critical component of the strategy for addressing the adverse effects of PM on visibility and on public welfare. Regional haze regulations under the authority of section 169A of the Act were deferred by EPA at the time the original visibility regulations were issued in 1980 due to the need for better technical tools and knowledge of the effect of fine particle constituents on visual air quality. Since 1980, technical tools have improved to support a regulatory program to address the long-documented regional haze visibility impairment. This finding is confirmed by various technical studies of haze, including the National Academy of Sciences 1993 study *Protecting Visibility in National Parks and Wilderness Areas* (NAS 1993, p. 11). In addition, section 169B of the Act calls for regulations to be developed under section 169A within 18 months of receipt by EPA of the recommendations from the Grand Canyon Visibility Transport Commission (the Commission's report was issued in June 1996). Thus, the new regional haze program will also make for a more comprehensive visibility protection program as required under section 169A.

(5) *Comment*: Some commenters stated that before a nationally applicable regional haze program can be established, additional visibility transport commissions should be formed for regions of the country other than the Colorado plateau, and certain critical databases and strategies should be developed.

Response: Since adopted in 1977, section 169A of the Clean Air Act has authorized EPA to address regional haze visibility impairment. Section 169A(a)(1) establishes as the national visibility protection goal "the prevention of any future, and the remedying of any existing, impairment of visibility in mandatory class I Federal areas which impairment results from manmade air pollution." Thus, the national goal provides for visibility protection generally. Further, the national goal is the lodestar for EPA's visibility protection regulations. In section 169A(a)(4) Congress delegated to EPA authority to issue regulations to assure "reasonable progress toward meeting the national goal." As explained in Maine v. Thomas, 874 F.2d 883, 885 (1st Cir. 1989, "EPA's mandate to control the vexing problem of regional haze emanates directly" from these provisions of the Clean Air Act. In adopting section 169A, Congress evinced its intent to address impairment caused by "hazes" and the potential corresponding need to control a "variety of sources" and "regionally distributed sources." H.R. Rep. No. 294, 95th Cong., 1st Sess. at 204. While EPA deferred addressing regional haze in its original 1980 regulations it did so because of technical obstacles, not because of a limitation on its legal authority. 45 Fed. Reg. 80,084 (Dec. 2, 1980). Indeed, in the 1980 rule EPA expressed its intent to address regional haze in a future rulemaking under section 169A.

The provisions in section 169B of the Clean Air Act, adopted in 1990, grew out of Congress' frustration that EPA had not more expeditiously addressed regional haze under its section 169A delegated rulemaking authority. Congress authorized visibility transport commissions under section 169B and provided for regional haze-related research to facilitate EPA's development of regional haze regulations. Congress made it clear that it did not intend section 169B to impinge upon EPA's long-standing obligation to address regional haze visibility impairment. See 136 Cong. Rec. S2878 (daily ed. March 21, 1990) (statement of Sen. Adams) ("[t]he authority to establish visibility transport regions and commissions is a supplement to the administrators [sic] obligation under current law" and "[t]he Administrator may not delay requirements under section 169A because of the appointment of a commission for a region under section 169B"); id. at S2887 (statement of Sen. Wirth); see also 136 Cong. Rec. H12883 (daily ed. Oct. 26, 1990) (statement of Rep. Wyden) ("[n]either the original House language nor the Senate language adopted in conference repealed or lessened EPA's obligations under the 1977 law"). Thus, visibility transport commissions are a potential tool for, but not a prerequisite to, the development of regional haze regulations.

Some of the commenters also raised concerns about the adequacy of representative monitoring data to support a national regional haze regulation, and other concerns related to the development and implementation of the regional haze rules. As noted, the National Academy of Sciences has concluded that there is an adequate technical basis for a regional haze regulatory program. In any event, the regional haze rules are in the process of being developed in a separate rulemaking proceeding, and EPA has reached no final decision about their content. Consequently, such comments are beyond the limited scope of the final NAAQS decision here but are properly raised in the actual regional haze rulemaking.

(6) *Comment:* Hunton and Williams for Utility Air Regulatory Group (IV-D-2250) stated that the EPA did not demonstrate that adverse effects to the public welfare would still exist if PM concentrations were reduced to the levels of the proposed PM2.5 primary standards (e.g., 15 micrograms annual average). Therefore, a secondary standard for visibility is not appropriate.

Response: See discussion of this issue in the preamble to the final PM NAAQS regulation and in response to comment (2) above. In the preamble to the final PM NAAQS, EPA references quantitative information demonstrating how fine particle levels at and below the level of the PM2.5 primary standards result in known or anticipated adverse effects on visibility, and indicates that the Grand Canyon Visibility Transport Commission and Southern Appalachian Mountain Initiative have expended significant time and resources assessing adverse visibility impacts in their respective regions. EPA believes that these adverse effects below the level of the primary standard are best addressed through a secondary standard equivalent to the primary standard in combination with implementation of a regional haze program by the States.

(7) *Comment*: Several commenters stated they did not believe that the proposed combination of secondary standards equivalent to the primary PM standards and a regional haze program would protect or improve visibility in urban areas (particularly in the western U.S.), in rural areas, and in non-class I areas.

Response: See section III of the preamble to the final PM NAAQS regulation. EPA believes that due to the regional nature of the regional haze problem, strategies to improve visibility in the 156 mandatory class I areas under a regional haze program will also benefit non-Class I areas. For example, PM emission reductions may be needed in and around certain urban areas in the West in order to make reasonable progress toward the national visibility goal in nearby class I areas, even though some of these areas are not expected to exceed the $PM_{2.5}$ standards. As noted in the preamble, EPA also intends to pursue opportunities to obtain additional information on urban and non-Class I area visibility and other welfare effects in the future through examination of available fine particle monitoring data. Current or planned monitoring networks and initiatives, such as monitoring and chemical analysis of $PM_{2.5}$ in urban and rural background sites, efforts to better characterize real-time environmental conditions in major population centers, and new automated airport visibility monitoring networks should provide data needed to evaluate trends in these areas. This information should also help to better characterize the nature and spatial extent of urban and non-class I visibility problems and thus serve to inform future decisions on NAAQS revisions or other appropriate measures.

(8) *Comment*: Several commenters stated that establishing the proposed secondary NAAQS equivalent to the primary NAAQS would allow additional short-term episodes that would cause visibility impairment, public safety concerns, materials damage, soiling, and nuisance effects. One commenter suggested a short-term secondary standard substantially below 50 ug/m3 to protect against visibility effects and encouraged the

development of strategies beyond the primary NAAQS level.

Response: See section III of the preamble to the final rule. The annual and 24-hour secondary $PM_{2.5}$ standards will provide a significant, nationally uniform degree of protection against urban and broad regional visibility impairment while avoiding unduly harsh or burdensome impacts on any particular geographic area. The EPA acknowledges that this level of protection will not eliminate all localized instances of urban visibility impairment or adequately address impairment in western Class I Federal areas, due to various regional factors such as background PM levels, humidity levels, population density, and industrial activity, as well as local topography. In making its recommendations, the National Academy of Sciences report noted the difficulties presented by these regional variations in attempting to address all instances of visibility impairment, and therefore found that regional approaches would generally be more appropriate and responsive (NAS, 1993, p. 6). For this reason, in addition to the uniform degree of protection accorded by the secondary standards, welfare effects caused by impairment of visibility in mandatory Class I Federal areas will be addressed through a regional haze program under section 169A of the Act. A regional haze program will provide States the ability to address such nationally significant but regionally variable welfare effects through appropriate regional approaches. As noted in previous responses to comments, to the extent that such approaches are needed to address visibility impairment in a number of contiguous or geographically proximate Class I areas, significant visibility and other benefits are expected to occur in nearby urban areas as well. As noted in the preamble, EPA believes that urban visibility impairment that is of concern in particular urban areas, but is not addressed through these regional approaches, is most appropriately addressed through State or local initiatives, as has been done through adoption of a local visibility standard for the city of Denver, Colorado.

2. Specific scientific/technical comments.

(1) *Comment* (Electric Power Research Institute, PM, IV-D-2329): The statement that one deciview represents the threshold of humanly perceptible change in visibility is an overstatement.

Response: In the December 1996 PM NAAQS proposal (61 FR 65663), EPA stated in a footnote that "[U]nder many scenic conditions, a change of 1 deciview is considered perceptible by the average person." EPA did not state that this was the "threshold of humanly perceptible change."

Due to the broad variety of scenic, atmospheric, and lighting conditions at mandatory Class I Federal areas across the country, at any specific time a given area may contain vistas for which slightly more or less than one deciview above background conditions represents a perceptible impact for the components of the scene. For example, a view of a snow-capped mountain may be more sensitive to changes in air quality than a view of a forest with the result that less than a 1.0 deciview change is perceptible for that portion of

the scene. Conversely, in another scene a deciview change slightly greater than 1.0 may not be perceptible. In general, however, EPA believes that a 1.0 deciview change represents a perceptible change across the range of complex views found in all Class I areas. It is for this reason that in Appendix G to the PM Staff Paper, EPA stated that the deciview metric "may be useful in defining goals for perceptible changes in visibility conditions under future regulatory programs."

(2) *Comment* (National Mining Association, PM, IV-D-2158): Coarse particles at current ambient levels present no substantial health or welfare concerns.

Response: As discussed in section II above, EPA disagrees with this comment with respect to the effects of coarse PM on health; in addition, EPA disagrees that coarse PM has no effect on visibility or soiling and nuisance. Although fine particles are often the principal cause of visibility impairment in rural areas, coarse particles can account for a significant fraction of the light extinction budget at many western locations. Based on a review of 1993-95 data from the IMPROVE visibility monitoring network, coarse particles account for 10-23% of annual average reconstructed light extinction in many parts of the western U.S. outside the Pacific Northwest. Locally high concentrations of coarse particles can appear as visible dust clouds and "plume blight," and in such locations, coarse particles contribute significantly to soiling and nuisance effects.

(3) *Comment*: One commenter (IV-D-2329, Electric Power Research Institute) stated that the levels describing background PM and visibility conditions in the 1990 National Acid Precipitation Assessment Program study (Report 24, Visibility: Existing and Historical Conditions - Causes and Effects) were not endorsed by the scientific community.

Response: EPA disagrees with the commenter. The NAPAP study was subject to extensive peer review and the background values in the visibility chapter subsequently have been referenced in numerous other publications. Furthermore, the assessment of background levels included in the final Criteria Document and Staff Paper was the subject of significant discussion and review by CASAC. Although subject to recognized uncertainties, these estimates were peer reviewed and reflect the best available judgment on background based on the available scientific information.

C. Federal Reference Method for Monitoring $PM_{2.5}$

EPA proposed a new $PM_{2.5}$ reference method to determine attainment of the standards. EPA proposed to base the new method on a conventional type filter sampler that collects 24-hour integrated $PM_{2.5}$ samples on a 47 mm Teflon filter that is subsequently moisture and temperature conditioned and analyzed gravimetrically. The sampler is a low volume sampler that operates at a flow rate of 1 cubic meter per hour, for a total sample volume of 24 m^3 for the specified 24-hour sample collection period. The sampler is relatively modest in cost and easy to operate, operates over a wide range of ambient conditions, produces a measurement that is comparable to

large sets of previously collected PM data in existing databases, and provides a physical sample that can be further analyzed for chemical composition. Numerous commenters provided technical comments on the design and operation of the instrument, quality assurance and specifications, as well as more general comments on the indicator and the suitability of the proposed method for measuring components of $PM_{2.5}$.

1. General Comments on Sampler Design

This section addresses comments in category I.B of the PM Monitoring Support Summary of Public Comments Document.

(1) *Comment*: Commenters suggested the use of a different indicator, use of a different size cut, inclusion of additional, more specific, constituents (e.g., acid aerosols, carbon, metals, and semi-volatiles), and/or use of a multi-filter method.

Response: Comments on the specific indicator and size cut are discussed in section II above. Early in the development process these issues were extensively evaluated, and design decisions were based on this internal evaluation, public input, and the advice of CASAC, including a technical subcommittee on $PM_{2.5}$ monitoring. Other factors affecting the basic design of the method were the need for historical continuity, high measurement precision, and simplicity of operation, all in response to current national monitoring objectives and available time constraints imposed by the court ordered deadline. In selecting the basic measurement approach, substantial weight was given to maintaining comparability to $PM_{2.5}$ samplers, such as the "dichotomous sampler," that were used to obtain much of the data upon which the new standards are based. After evaluating these issues in light of the comments received, EPA concludes that the conventional PM measurement approach best meets the objectives and will provide $PM_{2.5}$ measurements that are comparable to the air quality data used in the health studies that provide the basis for the $PM_{2.5}$ standards.

Although the sampler is conventional in configuration, its design is much more sophisticated than that of previous PM samplers. This more sophisticated sampler, together with improved manufacturing and operational quality assurance, is necessary to achieve the more stringent data quality objectives established for $PM_{2.5}$ monitoring data. To meet precision requirements, the critical mechanical components of the inlet, particle size separator, downtube, and upper portion of the filter holder are specified by design, as proposed. Almost all other aspects of the sampler are described by performance-based specifications, also as proposed.

Although, for the reasons specified above, the FRM design and characteristics has focused on measurement $PM_{2.5}$ for attainment determinations, EPA strongly encourages the development and use of additional instruments capable of measuring specific substances and fractions of PM. Such information can be invaluable in identifying sources of various components for implementation as well as in conducting research on

health effects.

(2) *Comment:* Several commenters felt that the portions of the sampler that were specified by design would stifle further improvements and innovations.

Response: Although the EPA specifies methods by performance whenever possible, for the $PM_{2.5}$ reference method, development of adequate performance specifications for inlet aspiration and particle size discrimination would have been a very difficult, costly, lengthy, and problematic process. Moreover, manufacturer testing of proposed inlet and particle size discrimination devices against such performance specifications would require elaborate specialized facilities and would be extremely costly. For these reasons, the EPA believes that specification of these critical components by design is a prudent and very cost-effective way to ensure good inter-manufacturer and intra-manufacturer precision of the $PM_{2.5}$ measurements. Therefore, these components are specified by design, and other aspects of the sampler are specified by performance, as proposed.

EPA encourages innovations and improved samplers or measurement methods and provided for as Class II and III equivalent methods (see 40 CFR Part 53), in particular the development of continuous or sequential samplers to facilitate more comprehensive sampling. As noted above, EPA encourages use of innovative and supplemental approaches for special purpose monitoring to support implementation and scientific research.

2. Inlet and Impactor Design

This section addresses comments in category I.B.3 and I.B.4 of the PM Monitoring Support Summary of Public Comments Document.

(1) *Comment:* Several commenters addressed the inlet design, noting that the inlet could allow entrance of precipitation and possibly insects.

Response: In fact, the PM_{10} inlet selected for the sampler has been used effectively for many years to obtain many of the $PM_{2.5}$ and PM_{10} measurements that formed the basis of the epidemiological studies. While EPA acknowledges that there have been some reports of intrusion of precipitation, the Agency believes the problem is relatively minor. Nevertheless, the inlet has been modified to further reduce the possibility of precipitation (and possibly small insects) reaching the sample filter to damage the $PM_{2.5}$ sample. Extensive wind tunnel tests have shown no significant compromise in the $PM_{2.5}$ aspiration performance of the modified inlet.

In addition, a new provision has been added, in section 7.3.8, to require that the sampling air entrance of the inlet be at a specific height of 2 ±0.2 meters above the supporting surface to help ensure homogeneous air samples when collocated samplers of different

types are operated simultaneously.

(2) *Comment:* Other commenters addressed the sharpness of the size cut and how it is obtained (e.g., whether more than two stages should be used and what size cut should be used for each stage).

Response: These aspects were carefully considered in selecting the sampler configuration. The selection by EPA of the previously used PM_{10} inlet established the size cut for the first stage, and the second stage was designed to be simple, reliable, and low in cost for user agencies. In EPA's estimation, the advantages of this configuration outweigh any modest advantage that might have been gained by designing a new inlet/separation configuration that would further refine the cut points at each of two (or more) stages.

(3) *Comment:* A few commenters questioned whether the inlet was wind speed dependent at high wind speeds.

Response: The selected inlet has been shown to perform well up to 24 km/hr with 10 μm aerosols and is expected to perform well at higher speeds with 2.5 μm aerosols. The EPA again determined that the advantages of using the selected inlet (modified to reduce precipitation intrusion) outweighed the possible minor improvement in wind-speed characteristics that might have been obtained in a newly-designed different inlet.

(4) *Comment:* Some commenters felt that other types of particle discrimination techniques such as cyclones, virtual impactors, etc. should be allowed.

Response: Again, these alternatives were evaluated and the specified inlet and impactor were determined to best meet the various objectives of the sampler. However, EPA has provided for consideration of other particle size selection techniques or devices for approval if incorporated into candidate equivalent methods for $PM_{2.5}$.

(5) *Comment:* Several commenters addressed the impactor design, noting that the impactor should be changed (1) to sharpen the size-cut characteristic, (2) to address concerns regarding possible contamination and/or performance loss due to impactor oil, and (3) to improve ease of access to service.

Response: To address the first concern, the initial prototype impactor presented to CASAC in February 1996 had been modified slightly to sharpen its size-cut characteristic. The current impactor is designed to lower cost and to optimize cut sharpness, loading capacity, manufacturing simplicity, manufacturing quality control, serviceability, and reliability. A report containing the penetration efficiency of the impactor is available in Docket A-95-54. With regard to impactor oil concerns, the impactor oil selected has a very low vapor pressure, and testing has indicated no contamination of the sample filters with impactor oil. The EPA believes that the

impactor design is as accessible as possible, given the design objectives. Some flexibility may be allowed for manufacturers to develop improved closure devices or other external modifications. Proper maintenance will, of course, be very important and will be stressed in the associated operator instruction manuals and in other training and guidance materials. The EPA has been performing field and laboratory tests that will provide the basis for detailed guidance for all necessary preventive maintenance. Proper installation procedures for the oil and the impactor filter, as well as all other maintenance requirements, will be available in the quality assurance procedures and guidance contained in a new section 2.12 to be added to the EPA's "Quality Assurance Handbook for Air Pollution Measurement Systems, Volume II, Ambient Air Specific Methods" (EPA-600/R-94/038b).

3. Anodized Aluminum Surface

All internal surfaces exposed to sample air prior to the filter are required to be anodized aluminum (section 7.3.7).

Comment: A few commenters expressed concern that the anodized aluminum surfaces in high volume PM_{10} samplers have shown substantial pitting, particularly in the venturi flow control device.

Response: The anodized aluminum surfaces are required in the $PM_{2.5}$ sampler to maintain comparability to previously used samplers. The EPA believes that the much lower flow rate in the $PM_{2.5}$ sampler will greatly reduce the pitting tendency, and the active flow control in the $PM_{2.5}$ sampler is not dependent on the physical dimensions of a critical orifice as it is in a venturi flow control device.

4. Filter for $PM_{2.5}$ Sample Collection

The proposed reference method called for the sample to be collected on a 47 mm Teflon filter. Many of the comments received on the measurement method concerned the proposed filter medium and its performance. This includes comments in category I.B.5 through I.B.8 of the PM Monitoring Support Summary of Public Comments Document.

(1) *Comment:* Commenters expressed concerns with the use of Teflon filters and with the selection of a single-filter method (as opposed to a multiple-filter technique). Several commenters recommended that alternative filter media be allowed, in most cases to support more detailed analysis and/or to allow the capture of additional PM components. Other comments noted potential advantages of other media in operating characteristics or chemistry requirements. Operational concerns expressed about Teflon filters included propensity for tearing, possible loss of integrity, and relatively high cost. Other concerns were that Teflon is generally not conducive to carbon analysis, and that Teflon filters may not hold deposited PM. Many commenters recommended use of a multi-filter sampler to support chemical speciation analysis in addition to the compliance

determination.

Response: To address some of these general concerns about the performance of the specified filter material, some minor refinements to the filter specifications concerning the filter diameter and the filter support ring have been made to ensure proper performance of the filter in the specified filter holder. Additional clarifications have been made to the maximum moisture pickup and the filter weight stability requirements. Although Teflon may preclude certain chemical analyses (e.g., elemental and organic carbon), Teflon was used to derive $PM_{2.5}$ mass in epidemiological studies and the EPA believes that Teflon filter material is the best overall choice to meet the objectives of compliance monitoring and to provide good measurement precision. Other filter media are likely to provide reduced gravimetric precision and preclude more types of subsequent chemical analysis. The regulations, do not preclude the use of alternative filter media for use in special purpose monitors. Furthermore, additional or alternative samplers or filter types can be considered as candidate equivalent methods and can be used for non-compliance monitoring, where necessary.

The EPA reiterates that compliance monitoring based on mass concentration of $PM_{2.5}$ is the primary objective of the reference method. Multi-filter capability would have substantially increased the cost and complexity of the sampler. However, multi-filter samplers can be considered as candidate equivalent methods. In addition, multi-filter samplers can be used as special purpose monitors (SPMs) to perform characterization studies, develop control strategies, and conduct other special studies as has been done previously for PM_{10}.

In response to numerous comments received on Appendix L and on the provisions of part 58 regarding the need for chemical speciation, the EPA is requiring a chemical speciation trends network and is assigning a high priority to additional chemical speciation through its 105 grant allocation program. EPA will issue guidance describing the monitoring methods and scenarios under which speciation should be performed. The program will incorporate additional $PM_{2.5}$ samplers that allow for the simultaneous collection of aerosols on multiple filter media.

(2) *Comment:* Although a few commenters generally supported the requirement for archiving filters, many questioned the provision. Among the concerns of these commenters were the economic burden of archiving filters for such a long time, whether the potential existed for archiving to introduce artifacts and potential bias, and whether the number of filters archived and length of time should be reduced.

Response: The associated requirement for archiving filters has been removed from Appendix L (section 10.17) and relocated to 40 CFR, part 58, Appendix A. The basis for this change is that filter archiving is a supplemental monitoring requirement and not an integral part of the reference method for determining compliance with the $PM_{2.5}$ NAAQS. Furthermore, in Appendix A, the length of time for archiving filters has been reduced to 1

year to respond to these concerns.

(3) *Comment:* Several commenters to 40 CFR part 58 expressed concern that the recommended sampling frequency for many sites would require the timely development of sequential samplers.

Response: As part of this effort to expedite the development of sequential filter samplers, appropriate provisions of Appendix L have been clarified to apply not only to a single-sample sampler, but also to a sequential-sample sampler, provided that all specifications are met and no deviations, modifications, or exceptions are made to the inlet, downtube, impactor, or the upper portion of the filter holder (all of which are specified by design). Samplers that have minor changes or modifications in these components, have changes that alter the aerosol's flow path, or contain other significant deviations will be required to meet the requirements of Class I equivalent methods (see amendments to part 53 elsewhere in the Federal Register). Further, a provision has been added to require that sequential sample filters stored in a sequential sampler be adequately covered and protected from contamination during the storage periods in the sampler.

(4) *Comment:* A few commenters expressed concern about who must carry out filter tests to determine if they meet the filter specifications.

Response: In response, the filter specifications have been clarified to indicate that filter manufacturers should generally carry out most or all of the filter performance tests in order to certify that their filters meet the filter specifications for the $PM_{2.5}$ reference method. In addition, the EPA has historically conducted acceptance tests on filters procured for NAMS/SLAMS networks prior to distribution to State and local agencies and plans to continue this practice for the new $PM_{2.5}$ sampling program.

(5) *Comment:* Some commenters requested additional information on the requirement that an ID number be attached to each filter.

Response: Preliminary information indicates that it is not practical at this time for either filter manufacturers or users to print an ID number directly on the filter. However, the EPA is continuing to pursue this goal. In the meantime, alternative means, such as attaching an appropriate ID number to the filter's storage container, will be necessary. Additional details and possible alternative filter identification methods will be provided in the new Section 2.12 of the "Quality Assurance Handbook for Air Pollution Measurement Systems, Volume II, Ambient Air Specific Methods."

5. Filter Handling/Weighing/Conditioning Requirements

This section addresses comments in category I.B.6 through I.B.8 of the PM Monitoring Support Summary of Public Comments Document.

(1) *Comment:* Although many commenters felt that requirements to control variations in these procedures were necessary to control an important aspect of measurement uncertainty, many other commenters felt that the filter handling requirements for collected $PM_{2.5}$ samples were too burdensome.

Response: EPA believes that handling of the exposed filter between retrieval from the sampler and commencement of the conditioning period is expected to be one of the most significant sources of $PM_{2.5}$ measurement variability. Thus, EPA concludes that specific requirements for this activity are necessary.

(2) *Comment:* Some commenters felt that the samples should be kept cold until analysis to prevent volatile losses.

Response: In response to this concern, the restriction on the maximum temperature exposure for collected samples has been reduced from 32 to 25° C, and a recommendation has been added for sampler operators to keep the samples as cool as practical between retrieval from the sampler and delivery to the conditioning environment. Further, the length of time permitted between retrieval of the filter and post-collection weighing is increased from 10 to 30 days, *provided* that the sample is maintained at 4° C or less between retrieval and the start of the conditioning period. The new section 2.12 of the "Quality Assurance Handbook for Air Pollution Measurement Systems, Volume II, Ambient Air Specific Methods" will provide guidance and techniques for keeping samples cool during this period and may suggest devices to document maximum temperature exposure of the sample.

(3) *Comment:*Commenters also requested additional specifications and guidance for field blanks.

Response: The EPA will provide additional clarification and detailed procedures and guidance regarding field blanks in the new section 2.12 of the "Quality Assurance Handbook for Air Pollution Measurement Systems, Volume II, Ambient Air Specific Methods."

(4) *Comment:* Some commenters felt that the filter weighing requirements were too restrictive. Other commenters had questions about operational and calibration procedures for the balances.

Response: Because filter weighing is one of the most significant sources of $PM_{2.5}$ measurement variability, specific requirements and restrictions are deemed necessary. However, in response to some of the concerns expressed, the proposed requirement that both pre- and post-weighings be carried out by the same analyst has been reduced to a non-mandatory recommendation. Detailed recommendations and guidance on filter weighing and balance calibration, based on information obtained in current field tests, will be provided in new section 2.12 of the "Quality Assurance Handbook for Air

Pollution Measurement Systems, Volume II, Ambient Air Specific Methods."

(5) *Comment:* Several commenters questioned the filter conditioning requirements, with some requesting a lower humidity range.

Response: Since humidity can profoundly affect the weight of the $PM_{2.5}$ on the filter, EPA maintains that filter conditioning requirements need to be tight to control measurement variability and to ensure satisfactory precision. But in response to at least one of the concerns, the filter conditioning humidity requirement has been changed to allow conditioning at a lower relative humidity commensurate with the humidity during sample collection. Filter conditioning is permitted at a relative humidity within ±5 RH percent of the mean ambient humidity during sampling (down to a minimum of 20 RH percent) for samples collected at average ambient humidities lower than 30 percent. The EPA will provide further details on filter conditioning controls in new section 2.12 of the "Quality Assurance Handbook for Air Pollution Measurement Systems, Volume II, Ambient Air Specific Methods."

6. Sampler Performance Requirements

Several commenters addressed sampler performance requirements, including (1) sampler flow control specifications, (2) filter temperature control, (3) sampler performance under extreme conditions, and (4) data reporting. This includes comments in category I.B.9-I.B.17 of the PM Monitoring Support Summary of Public Comments Document.

(1) *Comment:* Although a few commenters felt that requirements for flow control are necessary to ensure accurate size classification, several commenters thought the flow control specifications were too stringent and should be modified.

Response: In response to concerns that various sampler flow control specifications are too tight, EPA contends that good flow control is necessary to maintain uniform sampling, to ensure correct particle size discrimination, and to control measurement variability. Sampler manufacturers have been able to meet the specified flow control requirements, and field studies to date confirm that prototype samplers are able to meet these flow control requirements.

(2) *Comment:* Several commenters questioned the requirement for the temperature of the filter to be maintained at or below ambient temperature plus 3°C. Some felt it was not needed if, after sampling, the filter could sit for several days at temperatures up to 32°C. Others questioned whether and how such tight temperature control could be achieved.

Response: In response to comments about the ambient temperature plus 3 °C filter temperature control requirement, EPA believes that fairly tight control of the sample filter temperature is necessary to minimize losses of semi-volatile components over a wide temperature range, and tight temperature control has been strongly recommended

by the Clean Air Scientific Advisory Committee (CASAC). Monitoring of the filter temperature difference from ambient temperature is necessary to verify that the sampler filter temperature control is functioning properly. Testing to date indicates that the proposed 3 °C (above ambient temperature) limit is somewhat difficult to meet; however, a 5 °C limit can be reasonably met. Therefore, the filter temperature control requirement has been relaxed slightly from 3 °C to not more than 5 °C above the concurrent ambient temperature. Ambient and filter temperature sensors will require periodic calibration or verification of accuracy. In response to a common comment, the method has been clarified to indicate that exceedance of the filter temperature difference limit would not *necessarily* invalidate the sample.

(3) *Comment:* Several commenters addressed the issue of sampler performance under extreme weather conditions (e.g., high or low temperatures, low pressures, high winds, high or low humidity, fog, and dust storms). Some commenters wanted the operating and testing conditions to be expanded. Other commenters expressed concern about reduced data capture implications if samplers fail to operate under extreme conditions. Other commenters were concerned with the burden to State and local agencies to get the samplers to work under such conditions.

Response: In response to concerns about the performance of the sampler under extreme weather conditions, the EPA has established sampler specifications that are intended to cover reasonably normal environmental conditions at about 95 percent of expected monitoring sites. The sampler qualification tests in 40 CFR part 53 address most, if not all, of these operational requirements. Specification of the sampler performance for sites with extreme environmental conditions would substantially raise the cost of the sampler for other users, most of whom do not require the extra capability. Users requiring operation of samplers under extreme conditions are encouraged to develop supplemental specifications for modified samplers to cover those specific conditions. Sampler manufacturers have indicated a commitment to respond to the need for modified samplers for such extreme conditions.

(4) *Comment:* Several commenters expressed concerns about data reporting. Many questioned the need to report so much information, and some were concerned about the effort to develop data management systems to comply with the reporting requirements.

Response: Although concerns were expressed that the amount of data required to be reported from each sampler is excessive, EPA stresses that only a portion of the data collected by the sampler needs to be reported to AIRS. These limited data requirements (i.e., ambient and filter temperature, barometric pressure, sample volume, variation in sample flow rate) are important to establish or verify the reliability and confidence of the $PM_{2.5}$ measurements and to aid in utilization of those data. The substantial additional data are provided by the sampler for the site operator's use, to provide confirmation of a given sample's validity, and to aid in troubleshooting should outlier measurements appear in the monitoring data. A variety of current electronic devices and systems may

be used to acquire and handle the data, and these devices can easily accommodate the amount of data required to be reported, as well as the additional, optional data. Printers, modem connections, and alternative data output connections or devices are not precluded. A provision has been added in section 7.4.17 to require sampler manufacturers to make available computer software to that will translate sampler output data into a standard spreadsheet format (since no specific format is specified for output of the sample data acquired by the sampler).

III. RESPONSES TO COMMENTS ON LEGAL, ADMINISTRATIVE, AND PROCEDURAL ISSUES RELATED TO THE REVIEW OF THE PM NAAQS

These responses address comments summarized in section IV and elsewhere in the Summary of Comments. Because of the emphasis public commenters placed on certain issues, EPA responded directly to them in sections IV and VIII of the preamble to the final rule. Section IV of the preamble addresses the following legal and procedural issues: (1) whether EPA must consider costs and similar factors in setting NAAQS; (2) whether EPA erred in its selection of a methodology for determining the level of a NAAQS that protects public health with an adequate margin of safety; (3) whether EPA committed a procedural error by not entering into the rulemaking docket underlying data from certain epidemiological studies; and (4) whether the 1990 amendments to the Act preclude EPA from revising the PM NAAQS to establish a new $PM_{2.5}$ indicator. Section VIII of the preamble addresses issues raised in public comments with respect to EPA's obligations under the Regulatory Flexibility Act, as amended by the Small Business Regulatory Enforcement Fairness Act (SBREFA), and the Unfunded Mandates Reform Act.

A. General Issues

1. Cost consideration

This section addresses comments that EPA erred in not considering costs or other societal impacts in establishing NAAQS, particularly for non-threshold pollutants or in light of uncertain health effects information. Some commenters also maintained that costs should be considered in setting secondary NAAQS.

(1) *Comment*: EPA is not precluded from considering costs and similar factors; among other things, the judicial decisions relied upon by EPA as precluding consideration of such factors rest on faulty analysis that predates and cannot survive scrutiny under Chevron, U.S.A. v. Natural Resources Defense Council, 467 U.S. 837 (1984) (AISI, IV-D-2242; GM, IV-D-2694; ATA, IV-D-2245; UARG, IV-D-2250).

Response: Post-Chevron decisions have confirmed that costs and similar factors may not be considered in setting NAAQS. See section IV.A of the preamble to the final rule.

(2) *Comment*: Section 109 of the Act does not preclude consideration of cost/benefit analysis

when read in pari materia with sections 108(a) and 302(h) (GM, IV-D-2694).

Response: See section IV.A of the preamble to the final rule. Reading section 109(b) together with sections 108(a) and 302(h) does not alter the conclusion that consideration of costs and similar factors is precluded in setting NAAQS. Section 109(b) provides that NAAQS are to be based on air quality criteria issued under section 108(a)(2). As the commenter indicates, section 108(a)(2) provides that information on welfare effects, as well as health effects, is to be included in the air quality criteria. That information, however, does not include costs and similar factors resulting from efforts to attain or maintain the NAAQS. Although section 302(h) defines "welfare" as including "effects on economic values," this phrase refers to the economic costs or pollution, not to the costs of controlling pollution. Lead Industries Ass'n v. EPA, 647 F.2d 1130, 1148 n.36 (D.C. Cir. 1980). Cf. Natural Resources Defense Council v. EPA, 824 F.2d 1146, 1157-59 (D.C. Cir. 1987).

(3) *Comment*: Section 109(d)(2)(c)(iv) requires EPA to consider the advice of its independent science advisors on any "adverse public health, welfare, social, economic, or energy effects" that might arise from implementing revised standards when establishing them (UARG, IV-D-2250).

Response: See section IV.A of the preamble to the final rule.

(4) *Comment*: If Congress intended to forbid consideration of costs and benefits under section 109, it would have enacted a preclusive section 302 definition of "health effect" or "margin of safety" similar to the section 302(h) definition of welfare effect (GM, IV-D-2694).

Response: Such a definition was unnecessary in view of other indicia of congressional intent discussed in section IV.A of the preamble to the final rule. See, e.g., Natural Resources Defense Council v. EPA, 824 F.2d 1146, 1157-59 (D.C. Cir. 1987).

(5) *Comment*: The selection of the level(s) for the PM NAAQS is a policy decision that must be cost-effective. If cost-effectiveness is not considered, there is no justification for stopping short of a "no adverse effects" level (AAMA, IV-D-2243).

Response: See sections IV.A and IV.B of the preamble to the final rule. Where there is a continuum of health effects for a given pollutant, with no apparent threshold, there is no indication that Congress intended the Administrator to eliminate all risks, and some indication that it did not. See e.g., H.R. Rep. No. 95-294, at 127 (1977). See also, e.g., Natural Resources Defense Council v. EPA, 824 F.2d 1146, 1153 (D.C. Cir. 1987) ("'safe'" does not mean "'risk-free'"). In such cases, section 109(b) essentially requires the Administrator to select a standard which, in her judgment, based on consideration of the nature and severity of the health effects involved, the size of the sensitive population at risk, and other health-related factors, will reduce the risks sufficiently to avoid

unacceptable risks. Consideration of costs is not necessary to such a determination. See id. at 1165.

(6) *Comment*: Consideration of costs is especially appropriate in the context of a decision to regulate a new pollutant ($PM_{2.5}$) for the first time (GM, IV-D-2694).

Response: See section IV.A of the preamble to the final rule. In establishing $PM_{2.5}$ as an additional indicator for PM, EPA is not creating a newly regulated pollutant within the meaning of the Act. The regulated pollutant has been and continues to be PM.

(7) *Comment*: The language and legislative history of the CAA do not require EPA to ignore practical consequences when establishing a $PM_{2.5}$ standard based on ambiguous health data (AISI, IV-D-2242).

Response: See section IV.A of the preamble to the final rule.

(8) *Comment*: When construed together with E.O. 12866, UMRA and SBREFA, the CAA allows EPA to consider the economic consequences of a new $PM_{2.5}$ standard (AISI, IV-D-2242).

Response: See section IV.A of the preamble to the final rule.

(9) *Comment*: The EPA's evaluation of health effects in connection with the proposed $PM_{2.5}$ standards is flawed because it fails to consider the direct relationship between income and mortality. The cost of partial attainment equates to approximately 400 statistical lives lost (AAMA, IV-D-2243).

Response: See section IV.A of the preamble to the final rule. NAAQS are to be based on health effects caused by pollution, not effects that might result from control of pollution. Natural Resources Defense Council v. EPA, 902 F.2d 962, 972-73 (D.C. Cir. 1990).

(10) *Comment*: The Act prohibits consideration of non-health matters in setting standards; the place to consider costs is in the development and adoption of State implementation plans (NRDC, IV-D-2244).

Response: EPA agrees.

(11) *Comment*: In selecting among alternative secondary standards that provide a safe environment, the purposes of the Act (section 101(b)(1)) require EPA to conduct a holistic inquiry into all effects on the public welfare to ensure that its standard-setting will actually advance the public welfare. In doing so, EPA must take into account adverse social and economic effects that might result from implementing a secondary standard, as evidenced by the requirement (section 109(d)(2)(c)(iv)) that CASAC advise EPA on such effects (UARG, IV-D-2250).

Response: The comment is flawed in several respects. First, it appears to assume that proposal of alternative standards amounts to a finding that any of the alternatives would

provide adequate protection of public welfare under section 109(b). This is incorrect; proposal of alternative standards (primary or secondary) reflects the Administrator's awareness that there may be a range of views on the scientific information on which NAAQS are to be based, as well as on how the information should be used in making the policy judgments required for the final choice of a standard. Only the Administrator's final decision on a standard, taking into account public comments on the proposal, can be said to represent her determination of what standard meets the statutory criteria.

Second, the purposes of the Act are stated in general terms and are, at best, only a general guide to decisions under specific sections of the Act. To the extent they appear to conflict with more specific decision criteria stated in the statutory provision at issue, the more specific criteria are controlling. Under the decision criteria stated in section 109(b), consideration of costs and similar factors is precluded in setting NAAQS, and CASAC's responsibility to advise EPA on adverse effects that might result from implementation of standards does not alter that conclusion. See section IV.A of the preamble to the final rule.

2. Margin of safety

This section addresses comments on the approach used by the Administrator in specifying a PM standard that protects public health with an adequate margin of safety.

(1) *Comment*: In setting a NAAQS with an adequate margin of safety, EPA must define what constitutes "acceptable risk" for present and future rulemakings. In reaching such a determination, EPA must consider among other factors the results of cost-benefit analyses, the acceptability of risk judged in a "real world" context, and any adverse public health effects that might result from implementation of alternative standards. In other words, EPA must adopt a specific approach for specifying a standard that protects public health with an adequate margin of safety, and that approach must consider costs and other societal impacts (UARG, IV-D-2250; GM, IV-D-2694; ATA, IV-D-2245).

Response: See sections IV.A and IV.B of the preamble to the final rule.

(2) *Comment*: In setting a NAAQS with an adequate margin of safety, EPA must first identify the lowest observed effects level and then apply a margin of safety to address uncertainties and to protect the most sensitive individuals within the at-risk population(s). The use of risk assessment in establishing NAAQS is a departure from past practice, and this departure was not adequately explained (Yuhnke, IV-D-2095).

Response: See Section IV.B of the preamble to the final rule and the further response in section II above on level.

3. 1990 CAA amendments

This section addresses public comments expressing the view that the 1990 amendments of the Act preclude EPA from adopting $PM_{2.5}$ as an additional indicator for PM and establishing standards for $PM_{2.5}$.

(1) *Comment*: The plain language of title 1, part D, subpart 4 of the 1990 amendments precludes EPA from regulating $PM_{2.5}$ (GM, IV-D-2694; AAMA, IV-D-2243).

Response: See section IV.D of the preamble to the final rule.

4. Data availability for key health studies

This section addresses comments that EPA erred by not obtaining and making publicly available certain raw "data" underlying key health studies.

(1) *Comment*: EPA improperly relied on factual data not available to the public, in that the underlying raw data for the Dockery et al. (1993), Pope et al. (1995), and Schwartz et al. (1996) studies was not made publicly available or placed in the docket (NMA, IV-D-2158; AAMA, IV-D-2243; UARG, IV-D-2250; AISI, IV-D-2242; API, IV-D-2247; GM, IV-D-2694).

Response: See section IV.C of the preamble to the final rule. See also section II.B of the preamble to the final rule and section II.A of this response-to-comments document.

(2) *Comment*: A number of commenters argued that EPA has the ability to obtain the data underlying the Dockery and Pope studies and that EPA has refused to exercise its authority to obtain this data (AISI, IV-D-2242; AAMA, IV-D-2243; NMA, IV-D-2158).

Response: See section IV.C of the preamble to the final rule. It is uncertain whether EPA does in fact have the legal authority to require access to the data. The tapes are not "subject data" pursuant to 40 CFR 30. 1130, Appendix C, since they were not developed, produced or generated with EPA funds. Although a legal argument may potentially exist that EPA has a right of access to this data, EPA has never previously asserted such an argument and the argument remains untested in the courts. EPA has not attempted to shield itself from this data and in fact, consistent with the legal rights and obligations of the researchers, EPA has urged that the data be made public. Pursuant to EPA's request, Harvard and HEI have agreed to re-examine the data and make the results of that review public. EPA's ability to rely on these studies without obtaining the raw data should not depend upon whether some agency of the federal government funded the science or on the extent of EPA's ability to access the data using extraordinary means such as those described above. As noted in the preamble, EPA did not rely upon the raw data underlying the Dockery and Pope studies; it relied upon the studies themselves and has disclosed these studies to the public for review and comment. Only in the most extreme cases--for example where there are credible allegations of fraud, abuse or misconduct

would a review of raw data be warranted. In this case, there is no such allegation and the studies in question have been subject to an extensive peer review process that has confirmed the scientific integrity of these studies and their suitability for use in the PM rulemaking. Moreover, as evidenced in the discussion of the selection of the standard levels in section II.F of the preamble to the final rule, even if these studies had not been considered in the final decision EPA would have selected the same annual $PM_{2.5}$ standard based on other short- and long-term exposure studies. For all of the above reasons, EPA does not agree that review of the underlying data for these studies is necessary. Furthermore, it remains unclear whether EPA has a legal right of access to such data and, if access is available, a legal right to distribute the data.

5. Availability of FRM test data for public review

This section addresses public comments that EPA erred by not entering into the docket certain field test data concerning the proposed Federal Reference Method for measuring $PM_{2.5}$ in the ambient air.

Comment: Proposal of the $PM_{2.5}$ Federal Reference Method was premature because field test data were not available in the docket, precluding meaningful public comment. EPA's failure to include these data in the record violates the requirements of section 307(d) of the Act and the Administrative Procedure Act (UARG, IV-D-2250).

Response: See section VI.B of the preamble to the final rule. See also the accompanying Federal Register notice announcing a supplemental comment period to take comments on field and other test data.

6. Consideration of disbenefits

Comment: One commenter (GM, IV-D-2694) theorized that reductions in fine particle levels resulting from the proposed $PM_{2.5}$ standards may result in adverse effects from increased UV-B radiation such as skin cancer, cataracts, and immunosuppression. A similar comment was made on the proposed O_3 standard. A closely related issue raised by the same commenter is that the presence of fine particles may also play a major role in counteracting global warming. Based on this possibility, commenter argued that global warming caused by greenhouse gases would be much more intense if aerosol levels were significantly reduced by the proposed fine particle standard. In essence, the commenter argued that EPA erred in not considering these risk-risk tradeoffs involved in revising the PM standards

Response: EPA strongly disagrees with this commenter's suggestion that such "disbenefits" of tighter standards can and should be considered in reviewing and revising NAAQS, because it is inconsistent with the Clean Air Act and ill-advised from an environmental management policy perspective. Furthermore, the commenter has not pointed to any quantitative assessment or scientific evidence that supports its claim that

the effects of implementing the final PM NAAQS would produce significant disbenefits. Each of EPA's reasons is discussed more fully below.

The clear intent of Congress in enacting the NAAQS provisions of the Clean Air Act prohibits EPA from considering in this rulemaking any health "disbenefits" that may result from the implementation of a new, more stringent NAAQS. Where the intent of Congress on a specific issue is clear, as determined by traditional tools of statutory construction, it must be given effect by the implementing agency and the courts. Chevron, U.S.A. v. Natural Resources Defense Council, 467 U.S. 837, 842-45 (1984). As described below, Congress clearly intended to limit EPA's consideration in developing criteria and in setting NAAQS to the adverse health effects caused by the presence in the ambient air of the pollutant in question. Accordingly, EPA is not considering in this rulemaking the alleged health "disbenefits" from implementation that have been raised by commenters, and EPA did not include them in the discussion of the air quality criteria.

The NAAQS provisions of the Act are set forth in sections 108 and 109 and were first enacted in 1970. In that year, Congress set up a three-step process for the development of NAAQS -- first, EPA must prepare a list of air pollutants meeting certain requirements; second, EPA must develop criteria for the listed pollutants; and third, EPA must establish NAAQS for the pollutants based on the criteria. See 42 U.S.C. sections 7408, 7409. At each step, there is evidence that Congress intended the Agency to consider only the adverse health effects caused by the presence in the ambient air of the pollutant at issue.

As the initial step, Congress directed EPA in 1970 to list "each air pollutant - (A) which in his judgment has an adverse effect on public health or welfare; (B) the presence of which in the ambient air results from numerous or diverse mobile or stationary sources; and (C) for which . . . he plans to issue air quality criteria . . . ." 42 U.S.C. section 7408(a)(1). In paragraph (A), Congress expressly focused the entire NAAQS process on pollutants that have an adverse or harmful effect on public health.

In the second step, EPA must develop air quality criteria for each listed pollutant. Section 108(a)(2) states that the "criteria for an air pollutant shall accurately reflect the latest scientific knowledge useful in indicating the kind and extent of all identifiable effects on public health or welfare which may be expected from the presence of such pollutant in the ambient air." Read out of context, the phrase "all identifiable effects" might be deemed sufficiently broad to encompass any health effect, whether positive or negative. But the phrase can clearly be read as meaning only harmful effects, and it is only part of a larger body of statutory language that evidences Congress' intent with respect to the NAAQS. Other language in sections 108 and 109 indicates that Congress had harmful effects of a pollutant in mind when it directed EPA to examine "all identifiable effects." Indeed, the immediately following sentence in section 108(a)(2) specifies three factors that the Agency must include in the criteria, and two of those three

factors expressly direct EPA to focus on "adverse" effects on health and/or welfare.[12] Similarly, the listing process in section 108(a)(1)(A) in 1970 required the Administrator to list for criteria and NAAQS development each air pollutant "which in his judgment has an adverse effect on public health or welfare . . . ."[13] Together, these statutory excerpts (with the provisions of section 109, discussed below) evidence Congress' clear intent for EPA to focus on the harmful effects of a pollutant in developing the air quality criteria.

Also, the express language of section 108(a)(2) limits the scope of causality that[14] it is appropriate for EPA to consider. The language directs EPA to focus on "effects that may be expected from the presence of the pollutant in the ambient air." This language parallels that in the listing process, which directs EPA to list pollutants "the presence of which in the ambient air" results from numerous or diverse mobile or stationary sources (section 108(a)(1)(B)). In both provisions, Congress limited the causality consideration to the effects caused by the emitted pollutant's presence in the ambient air. There is no language to support the idea that Congress intended to focus on the indirect effects of implementation efforts to reduce pollution following the establishment of a NAAQS. Indeed, such considerations would be premature at this point in the process, when the Agency is focusing on the criteria that will form the basis for setting the NAAQS.

In the third and final step, section 109 directs EPA to set the NAAQS based on the air quality criteria issued under subsection 108(a)(2). 42 U.S.C. 7409(b)(1)-(2). The case law on considering cost in NAAQS reviews confirms that Congress limited the Agency's consideration to the factors specified in section 108(a)(2). See section IV.A of the preamble to the final rule. Further, the 1970 Senate report evinces Congress' intent to

---

[12]The three factors are "(A) those variable factors (including atmospheric conditions) which of themselves or in combination with other factors may alter the effects on public health or welfare of such air pollutant; (B) the types of air pollutants which, when present in the atmosphere, may interact with such pollutant to produce an adverse effect on public health or welfare; and (C) any known or anticipated adverse effects on welfare." 42 U.S.C. 7408(a)(2)(A)-(C).

[13] In 1977, Congress amended the language in subsection 108(a)(1)(A). As revised, the subsection directs the Administrator to list each air pollutant "(A) emissions of which, in his judgment, cause or contribute to air pollution which may reasonably be anticipated to endanger public health or welfare. . . ." The legislative history shows that Congress inserted this revised language into the Clean Air Act in several sections to clarify that proof of actual harm was not necessary under section 108 or the other revised provisions of the Act, and to create a uniform test for regulation to protect public health and welfare. See, e.g., H.R. Rep. 95-294, at 43-51 (1977). The statutory language ("endanger") and the legislative history make it clear that Congress remained focused on the adverse effects of pollution. See id.

[14]In 1977, Congress also added a provision to address stratospheric ozone depletion and the increase in UVb radiation exposure that it causes. P.L. 95-95, sections 150-159 (1977).

focus on adverse health effects when setting primary standards. The report emphasizes that the Agency should protect the health of particularly sensitive citizens such as asthmatics, and declares that a NAAQS will be sufficient to protect the health of sensitive individuals "whenever there is *an absence of adverse effect* on the health of" an appropriate sample of such persons. S. Rep. No. 91-1196, at 10 (1970) (emphasis added).

Thus, it is clear from the language and legislative history of the 1970 amendments alone that Congress intended to limit EPA's focus to the adverse effects of a pollutant's presence in the ambient air. The repeated references to "adverse" effects, and Congress' focus on the effects caused by an emitted pollutant's presence in the ambient air, indicate that Congress did not want EPA to weigh the potential health "disbenefits" of pollution control against the adverse health effects from a pollutant's presence in the ambient air.

The 1977 and 1990 amendments to the Clean Air Act offer additional evidence confirming this conclusion.[15] In 1977 Congress made some significant changes to sections 108 and 109 but did not change its substantive instructions for setting NAAQS by amending subsections 108(a)(2) or 109(b). In new subsection 109(d), Congress directed EPA to review existing NAAQS periodically and established CASAC as a special advisory committee to advise the Administrator in such reviews. Congress expressly directed that both EPA's decisions and CASAC's recommendations on revisions of existing NAAQS be made in accordance with existing section 108 and subsection 109(b). 42 U.S.C. 7409(d)(1), (2)(A)-(B). As a separate task, Congress directed CASAC to offer advice to the Administrator in several areas, including any "adverse public health . . . effects which may result from various strategies for attainment and maintenance of such national ambient air quality standards." 42 U.S.C. 7409(d)(2)(C). This language specifically addresses the potential for health "disbenefits" from implementation. It shows that Congress was aware of the potential for such effects, yet declined to include them among the section 108 factors to be considered in setting a NAAQS. Instead, Congress directed CASAC to offer advice on the potential health effects of implementation separately from its involvement in the establishment and revision of the NAAQS. The legislative history confirms that such advice was intended for the benefit of the States and Congress, who might wish to use it in developing implementation strategies or in fashioning future legislation. See H.R. Rep. No. 95-294, at 183 (1977).[16]

---

[15] Even if doubt were to remain about Congress' intent after review of the 1970, 1977, and 1990 amendments, EPA's longstanding interpretation of the statutory language is clearly reasonable, for the reasons discussed above. Moreover, EPA's interpretation is supported by the policy reasons set forth later in this response.

[16] In 1977, Congress also added provisions to address stratospheric ozone depletion and the increase in UV-B radiation exposure that it causes. P.L. 95-95, sections 150-159 (1977).

In 1990, Congress again amended the Clean Air Act substantially without changing the basis for setting NAAQS. At the same time, Congress expressly addressed the issues of stratospheric ozone depletion and global warming that are the proximate causes of the health effects raised by commenters. Congress enacted Title VI (sections 601-618) to address stratospheric ozone depletion[17] and directed EPA in section 602(e) to consider the global warming potential of potential substitutes for stratospheric ozone depleting substances. These provisions demonstrate that Congress was aware of the potential environmental hazards of stratospheric ozone depletion and global warming but chose to address them separately from the process for setting and revising NAAQS. At the same time, other amendments show that Congress was aware EPA might revise the then existing NAAQS. For example, section 172(a)(1) expressly contemplates that EPA may revise a NAAQS in effect at the time of enactment of the 1990 Amendments.

The D.C. Circuit's decision in the PM_{10} case further supports the conclusion that Congress did not intend EPA to consider the implications for global warming and UV-B exposure from implementing strategies to reduce ozone and PM in accordance with the new NAAQS. In that litigation, AISI argued that EPA should have considered the potential human health effects of unemployment that might result from implementing the PM_{10} NAAQS. EPA had interpreted the statute as prohibiting the agency from considering such potential health effects of implementation in setting or revising a NAAQS. The court upheld EPA's conclusion, quoting subsection 108(a)(2) and stating that "it is only *health effects relating to pollutants in the air* that EPA may consider." Natural Resources Defense Council v. Administrator ("PM_{10}"), 902 F.2d 962, 972-73 (D.C. Cir. 1990) (emphasis in original).

As pertinent here, the potential health effects of UV-B radiation or global warming, like the potential health effects of unemployment, would not result from air pollution but from the implementation of pollution control. Like the potential health effects of unemployment, the potential health effects of both global warming and UV-B increases would not be caused by the presence of the applicable pollutant in the ambient air. In each case, there is an independent, intervening cause (unemployment, stratospheric ozone depletion, atmospheric increases in greenhouse gases) of the potential harmful effect. In each case, the argument for considering the potential effect is that implementing a new, tighter standard would "cause" an increase in the effect, but in each case the effect is actually a result of the intervening cause. In other words, without these intervening causes, there would be no health "disbenefit" to implementing the new NAAQS. In all three circumstances, the fact that the potential "disbenefit" would result from implementing the new NAAQS, rather than from the presence of the relevant pollutant in the air, means that EPA is prohibited from considering such effects.

The scenarios suggested by the commenter do differ from the unemployment concerns

---

[17] Title VI replaced the provisions regarding stratospheric ozone depletion enacted in 1977. P.L. 101-549, section 601 (1990).

presented in the $\underline{PM_{10}}$ litigation in one respect, and that difference argues yet more strongly against EPA's consideration of such concerns. In the $\underline{PM_{10}}$ litigation, AISI alleged that pollution control efforts would cause unemployment, which in turn would cause the harm to public health. In contrast, there is no causal connection whatsoever between ozone or PM reduction and either the buildup of greenhouse gases or stratospheric ozone depletion. Both of the environmental hazards cited by the commenter (and the health effects they potentially cause) would occur whether or not efforts were made to control PM or tropospheric ozone. All that this commenter alleges is that PM and tropospheric ozone mitigate the harm to public health caused by the independent environmental hazards known as global warming and stratospheric ozone depletion. Nothing in the statute or its legislative history suggests that Congress intended EPA to set a less protective NAAQS because the pollutant of concern might mitigate the harmful health effects of a wholly independent, environmental hazard. Indeed, as discussed above, the 1977 amendments and their legislative history indicate, to the contrary, that Congress did not intend EPA to set less protective NAAQS even if CASAC advised that implementation of NAAQS might cause adverse public health effects. Further, Congress' directive to protect particularly sensitive populations such as asthmatics would be vitiated if EPA had to set a less protective NAAQS to account for the NAAQS pollutant's potential to mitigate a different type of harm caused by an independent environmental problem that may affect other members of the public.

Even if the law had been written in such a way as to permit consideration of these hypothesized disbenefits and if, as the commenter has not shown, the available science permitted some quantification of such effects, EPA believes that it would be bad public policy to place any weight on this issue in reaching a decision on the PM standards. EPA does not believe it is appropriate or, as noted above, consistent with the intent of the framers of the Clean Air Act to consider increasing, or leaving at arbitrarily high levels, air pollutants that have direct effects on public health in certain sensitive populations in order to mitigate the effects of another pollution-induced problem, in this case increased UV-B or global warming. This would mean balancing the risks of adverse effects of breathing PM in the elderly, children that have asthma or other respiratory problems, and other sensitive groups with an attempt to intermittently reduce the risk of UV-B penetration that has been increased by CFC and other anthropogenic pollutants. Such a policy would ignore critical issues of equity and the distribution of relative risks.

Furthermore, this commenter has pointed to no convincing basis for considering such effects in the PM Criteria Document or elsewhere. The potential influence of aerosols in mid-latitude regions in controlling the transmission of solar UV-B radiation has been dealt with in only one article, Liu et al. (1991), which was mentioned by the commenter. This study was based on a model of the size distribution and optical properties of an idealized rural aerosol. Not enough information is available from this one calculation to permit an assessment of the effects of aerosol on UV-B transmission in urban areas or in rural areas with different aerosol properties. A quantitative risk assessment of the effects of aerosol reductions on the transmission of harmful UV-B radiation to the surface would

need to be based on radiative transfer calculations that incorporate data for the vertical distribution of the aerosol, the abundances of other scatterers and absorbers, the angular dependence of the scattered radiation (which requires information about the composition and size distribution of the aerosol), and the surface reflectivity (which requires information about surface characteristics) at a given location. This information is not available for potential changes in aerosol distribution that might result from meeting the fine particle standards. Moreover, because of the natural variability of pollutant concentrations in the atmosphere, the relative "protection" afforded by aerosol varies greatly from day to day, and any benefits would be irregular and unreliable. Estimates would have to be performed for a number of locations across the United States because of variations in all of the above parameters. In essence, the available information does not allow development of reliable estimates of any disbenefits.

Such information as is available suggests that any effect of a change in fine mass from current U.S. levels to those required by the new standards on aerosol UV-B shielding is likely to be small. Based on the observation that the stratosphere is responsible for on the order of 90% of shielding from UV-B, it is clear that the reductions in ozone depleting chemicals mandated by Title VI of the Act and the Montreal Protocol would dwarf the potential benefits of keeping tropospheric PM at levels that would not protect the health of populations sensitive to inhaling that pollutant. Reducing ozone depleting substances is clearly the appropriate way to address the effects of anthropogenic pollutants on UV-B, as it directly addresses the source of the largest problem.

Likewise, reliable assessments of the health effects due to climate change have yet to be performed. It is generally agreed that the negative radiative forcing exerted by tropospheric aerosols worldwide is sufficient in magnitude to offset the positive radiative forcing of the so-called greenhouse gases, at least in a globally averaged sense. However, it is not yet clear if this result can be directly translated into a cancellation of the potential climate response exerted by the greenhouse gases as the high degree of spatial heterogeneity of the aerosol burden may also lead to alterations in atmospheric dynamics. The tools to address these issues have yet to be developed.

For reasons detailed above, EPA rejects the recommendations of this commenter on legal, policy, and technical grounds. Most importantly EPA rejects the notion that air quality standards should be based on a principle of mitigating problems caused by anthropogenic emissions of pollutants such as CFCs and greenhouse gases by increasing -- or leaving at arbitrarily high values -- the levels of air quality standards for other pollutants whose presence in the air directly harms public health and welfare.

7. Miscellaneous comments

This section addresses various comments not addressed above.

(1) *Comment*: The EPA erred by not releasing, simultaneously with the revised air quality

criteria, guidance on control techniques as required by § 108(b)(1) of the Act (AAMA, IV-D-2243; GM, IV-D-2694; Exxon, IV-D-2113; NAIMA, IV-D-2161).

Response: Any such error would not affect the validity of the NAAQS themselves, which are to be based on air quality criteria containing the kinds of information specified in section 108(a)(2). In any event, section 108(b)(1) relates only to the initial issuance of criteria for a newly listed air pollutant. Where, as here, the Agency reissues such criteria for a NAAQS pollutant, the controlling provision is section 108(c). Section 108(c) states: "The Administrator shall from time to time review, and *as appropriate*, modify and reissue any criteria or information on control techniques issued pursuant to this section." 42 U.S.C. 7408(c) (emphasis added). As the statutory language makes clear, whether and when the modification or reissuance of a control techniques document is appropriate is left to the Administrator's discretion. In this instance, EPA has periodically issued control techniques information for specific particulate matter source categories such as residual wood combustion, agricultural and silvicultural open burning, and for sources of open fugitive dust as a means of supplementing the document Control Techniques for Particulate Emissions from Stationary Sources, Volumes 1 and 2, issued in September 1982. As part of EPA's continuing efforts to periodically update control techniques information, EPA plans to issue in the near future new guidance on control techniques for industrial stationary sources of particulate emissions to support the States and others in implementing the revised PM NAAQS including the $PM_{2.5}$ standards.

(2) *Comment*: EPA did not provide sufficient notice for the public hearings, and holding simultaneous hearings at four different locations and limiting the time for presentations to five minutes precluded effective participation; some commenters also sought additional hearings in the Southeast and the Southwest (ATA, IV-D-2245; BRPC, IV-D-7983; Russell Corp., IV-D-7990).

Response: When announcing the proposed decisions on November 27, 1996, EPA made widely available copies of the proposal notice, which clearly indicated that the date and location of the public hearing would be announced in a separate notice. Because of the strong public interest expressed, EPA decided to hold separate hearings at four locations to give interested parties more opportunity to participate. The EPA announced the dates, times, and locations of the hearings as soon as the necessary arrangements had been made - 3 weeks in advance. Because of the unusually large number of individuals who wanted to participate, it was necessary to limit oral presentations to five minutes. Under the circumstances, it was not feasible to hold public hearings in every region of the country; by holding four public hearings, EPA has more than satisfied the requirements of the Act to provide an opportunity for oral comments. In addition to the public hearings, EPA also solicited comment by voice mail, e-mail, fax, and written comments.

(3) *Comment*: EPA must explain the PM proposal's departure from prior decisions to reaffirm NAAQS for sulfur oxides and nitrogen dioxide, which were based on scientific evidence no more solid than the highly uncertain evidence on which EPA proposes a

revised PM standard (GM, IV-D-2694).

Response: The basis and rationale for EPA's decisions that revisions to sulfur oxides (SO_2) and nitrogen dioxide (NO_2) NAAQS were not appropriate were discussed in detail in the preambles to those rules (see 61 FR 25566, May 22, 1996; 61 FR 52852, October 8, 1996). Section II of the PM proposal notice (61 FR 65719, December 13, 1996) discusses in detail the basis and rationale for EPA's proposed decision to revise the PM standards. In EPA's periodic reviews of NAAQS, such factors as the nature and severity of the health effects involved, the size of sensitive population(s) at risk, the types of health information available, and the kind and degree of uncertainties that must be addressed vary from one pollutant to another. As a result, the decision whether and, if so, how the NAAQS for a given pollutant should be revised is necessarily specific to that pollutant and to the state of scientific knowledge available to the Administrator at the time of her decision. Thus, each standard review must be based on careful assessment of the available information in the air quality criteria for the pollutant in question. In the present case, EPA believes that the basis and rationale for the Administrator's decision to revise the PM NAAQS are fully explained in the preamble to the final rule and supporting documents, and that the decision is amply supported by the record.

(4) *Comment*. The EPA's decision to modify the NAAQS for particulate matter, which includes sulfates and nitrates derived from SO_2 and NO_2, but not the SO_2 and NO_2 standards, to control sulfates and nitrates appears inconsistent and arbitrary (GM, IV-D-2694).

Response: It has been EPA's longstanding position that secondary particles, such as sulfates and nitrates, should be considered during review of the PM NAAQS (see 49 FR 10408, March 20, 1984). The SO_2 and NO_2 NAAQS focus on the direct health effects of the gas phase precursors to PM.

(5) *Comment*: Revision of the PM standard is not appropriate or should be deferred in light of continuing improvements in air quality and reductions in PM resulting from the implementation of 1990 Act amendments (UARG, IV-D-2250; AAMA, IV-D-2243; AISI, IV-D-2242).

Response: Under section 109(d), the fact that air quality is improving is not an appropriate basis for declining to revise the NAAQS for PM or for deferring revisions the Administrator judges to be appropriate. Emission reductions to be achieved through the 1990 Amendments and resulting improvements in $PM_{2.5}$ ambient concentrations will be taken into account in developing control strategies for implementing the revised PM standards.

(6) *Comment*: EPA's proposal violates virtually every aspect of the legal standard for setting NAAQS under section 109, which requires EPA to demonstrate that the pollutant in question has an actual adverse effect on public health, because: 1) EPA has not

substantiated its methodologies for predicting health risk at levels below the current standards, 2) EPA has not established that the existing PM_{10} standards present an unacceptable risk of harm to public health, and 3) EPA has not shown that any $PM_{2.5}$ standard more stringent than the current PM_{10} standards is requisite to protect public health (UARG, IV-D-2250).

Response: Section 109(d) requires the Administrator to review the air quality criteria and NAAQS for a pollutant periodically and to "make such revisions in such criteria and standards and promulgate such new standards as may be appropriate" in accordance with sections 108(a) and 109(b), respectively. Section 109(b)(1) requires the Administrator to set NAAQS at levels which, in her judgment, will protect public health with an adequate margin of safety. Thus, where the Administrator finds that serious health effects occur, or may occur, on a widespread basis at pollutant concentrations lower than those specified in existing NAAQS, she clearly has discretion to conclude that revision of the NAAQS is appropriate. In doing so, she may weigh risks, project trends, extrapolate from limited data, and so forth, to carry out the preventive and precautionary purposes of the Act; proof of actual harm is not required. See, e.g., H.R. Rep. No. 95-294, at 43-51 (1977). See also, e.g., Lead Industries Ass'n v. EPA, 647 F.2d 1130, 1153-56 & n.50 (D.C. Cir. 1980). For discussion of the technical points raised by the commenter, see section II of the preamble to the final rule.

(7) *Comment*: EPA has not applied the proper legal criteria for selecting a new PM standard, because 1) EPA may establish a new NAAQS only if it first finds that the pollutant presents a "significant risk" to public health, and 2) EPA's proposal failed to focus on a representative sample of the sensitive population, as opposed to allowing the responses of particularly sensitive individuals within that group to drive the decision making process, thus departing improperly from past practice as in the SO_2 NAAQS decision (focusing on 20-25% of the sensitive individuals tested)(AAMA, IV-D-2243).

Response: Given the evidence that premature mortality and other serious health effects may occur at levels below the current NAAQS, the nature and potential magnitude of the public health risks involved, and the need to consider the fine and coarse fractions of PM_{10} as distinct classes of particles, both the Administrator and CASAC concluded that the current PM standards should be revised. This conclusion is amply supported by the record and was well within the Administrator's discretion under section 109(d). See section II of the preamble to the final rule and response to comment III.B.7(6) above.

The Administrator's decisions on the final PM standards are based on overall risks to the sensitive population and not on the responses of particularly sensitive individuals within that group. See section II.A of the preamble to the final rule. In any event, the Administrator's task in revising NAAQS is to select standards which, in her judgment, will protect the public health with an adequate margin of safety. The factors relevant to that determination vary from one pollutant to another, and NAAQS decisions are necessarily specific to the pollutant and to the record before the Administrator at the

time. See response to comment III.B.7(3) above. In other words, no single approach to determining what standards will protect public health with an adequate margin of safety is likely to be appropriate in all circumstances. Accordingly, both Congress and the courts have left to the Administrator's discretion the selection of an approach that will best fulfill the goals of the Act. See section IV.B of the preamble to the final rule.

(8) *Comment*: The public comment period on the PM proposal is inadequate and should be extended for periods ranging from 60 days to at least 1-year (Lange, Inc., IV-D-1257; Zurn, IV-D-5612).

Response: A 67-day comment period was originally provided, based on the schedule oredered by the court in American Lung Association v. Browner CIV-93-643 TUC-ACM (D. Ariz.). Additional time was available to interested parties because EPA distributed copies of the proposal notice widely when it was signed and announced on November 27, 1996, and the comment period itself did not commence until December 13, 1996. Also, the proposal was preceded by a lengthy scientific assessment process, in which the public had numerous opportunities, over a period of several years, to comment on EPA's assessments of the scientific information that was the basis for the proposal. At the request of commenters, EPA sought a 60-day extension of the public comment period, but the court only granted a 3-week extension.

(9) *Comment*: If rulemaking follows the current course, the final rulemaking will have to be vacated for failure to comply with section 307(d) because 1) underlying raw data has not been placed in the docket, 2) there is inadequate time for meaningful comment, and 3) there is inadequate time for EPA to respond to public comments. EPA has had numerous opportunities to discharge the court order, appeal it, or limit the impact of the order, and EPA has not taken advantage of those opportunities (UARG, IV-D-2250).

Response: With respect to the issue of underlying raw data, see section IV.C of the preamble to the final rule and response to comment III.B.5 (1) above. In EPA's judgment, the 89-day comment period, including the extension granted by the court, provided sufficient opportunity for the public to prepare and submit comments on the proposed rule. That this time was sufficient is evidenced by the number and volume of comments on these standards, which are more numerous and extensive than those for any prior NAAQS review. Finally, EPA believes there has been adequate time to respond to public comments as evidenced by the extensive discussion of comments in the preamble to the final rule and by responses to other significant comments in this document.

(10) *Comment*: EPA's use of the PM risk assessment exceeds the scope of EPA's authority under section 108 because EPA concedes it cannot establish whether the risk assessment is measuring effects of PM on public health or, instead, is measuring the effects of a complex mixture of pollutants in urban air for which PM may serve as an index. Although section 108(a)(2) allows some consideration of co-pollutants in establishing a NAAQS, EPA cannot rely exclusively and excessively on the measurement and effects

of co-pollutants in establishing NAAQS instead of evaluating the identifiable effects of PM. Further, EPA cannot rely on the risk assessment because the underlying "raw" data of the studies used as the basis for assessment were not publicly disclosed (NMA, IV-D-2158).

Response: Use of risk assessments per se is proper under section 109(b), in that NAAQS decisions are to be based on air quality criteria issued under section 108(a)(2), which are to include the latest scientific knowledge "useful in indicating the kind and extent" of the health effects that "may be expected" for varying ambient levels of the pollutant in question. The commenter's objection really goes to whether the pertinent epidemiological studies support the Administrator's conclusion that reported health effects are attributable to PM as opposed to a mixture of pollutants for which PM may serve as an index. As indicated in the Criteria Document (p. 13-31), reduction of PM exposure would lead to reductions in the frequency and severity of the health effects in question, whether the effects are attributable to PM, to a mixture of pollutants for which PM is a surrogate, or to both. In any event, EPA believes the Administrator's conclusion is amply supported by the record. See section II.B of the preamble to the final rule. Finally, the risk assessment, like the Administrator's decision on the final standards, is based on the pertinent epidemiological studies, not on the data underlying those studies, and the studies have been available for public scrutiny in the usual manner. In the circumstances presented, there is no bar to using the studies in the risk assessment or in the Administrator's ultimate decision on the standards. See section IV.C of the preamble to the final rule.

(11) *Comment*: The use of unpublished references in drafts of the criteria document and staff paper is inappropriate and unprecedented. EPA failed to make key references (e.g., Schwartz et al., 1996 manuscript) publicly available soon enough in the criteria document and staff paper development period to permit adequate time for public comment (API, IV-D-2247).

Response: It has been EPA's practice to cite papers that have been accepted for publication in the peer review literature, as well as other "gray" literature, during the preparation of drafts of the criteria document and staff paper so they can undergo the rigorous CASAC review process. Only those papers that are found to be of acceptable scientific quality are retained in the final documents. The EPA believes that there was sufficient time to review the Schwartz et al., 1996 manuscript during the development of the Criteria Document and Staff Paper. Interested parties also had opportunities to review and comment on the paper during the more than six month period between its acceptance for publication in a peer-reviewed journal, when it was made available as an "in-press" manuscript for citation in the Criteria Document, and the publication of the proposal and to submit comment on the published paper during the 89-day public comment period.

(12) *Comment:* The current review did not provide adequate time to fully assess the available

scientific information, particularly when contrasted to the last PM NAAQS review that took approximately 8 years to complete (AAMA, IV-D-2243, NMA, IV-D-2158).

Response: During the last review of the PM NAAQS, EPA developed three successive drafts of the joint Criteria Document for PM and sulfur oxides (SO_x), which added additional complexity to the task, for review by CASAC and the public. Interspersed with the preparation of these drafts, several workshops on different sections of the document were also held. The EPA also prepared two drafts of the PM Staff Paper that were reviewed by CASAC at two separate public meetings. Overall, the scientific assessment phase of the last review of the PM NAAQS, including CASAC's rendering of advice and recommendations for revised standards, was completed two years and three months after formal commencement of the review.

After formally commencing the present review, EPA held several workshops on key aspects of the Criteria Document and developed three successive drafts of all or portions of the document. The EPA also prepared two drafts of the Staff Paper. Throughout this process, the public had opportunity to express views at the public workshops, as well as at public CASAC meetings on the Criteria Document and Staff Paper drafts. Overall, the scientific assessment phase of the present review, including recommendations by CASAC and staff that the existing PM_{10} NAAQS be revised, was completed two years and two months after the initial announcement. This is only one month less than it took to reach a comparable point in the previous review.

While it is true that in the prior review it took an additional two years beyond CASAC closure on the science, to propose revisions to the original standards and an additional three years to promulgate those revisions, this is not an appropriate model for NAAQS reviews. The delay between completion of the scientific assessment phase and proposal was not occasioned by the need for further scientific assessment, but by the focus of EPA decision makers on unrelated issues, including a change in EPA management and the transition to a new Administrator and Assistant Administrator. Ultimately, these unrelated factors stretched the process to such a degree that, for both PM and SO_x, it was deemed necessary to update the Criteria Document and Staff Paper to reflect additional scientific findings. These updated findings did not alter the fundamental components of CASAC's recommendations or the proposed decisions on PM.

Although EPA has been under court order imposing a schedule for completion of the current review, EPA has sought and obtained modifications to provide additional time for the conduct of the scientific assessment phase of this review. As a result, the time provided for this aspect of the review was only two weeks shorter than the time EPA initially sought from the court. Notwithstanding the constraints imposed by the court order, EPA has conducted a thorough, comprehensive review of the scientific criteria and standards for PM. The procedures permitted full public participation in the process, and the time taken was commensurate with that taken in the previous review.

B. Regulatory and Environmental Impact Analyses

1. Compliance with E.O. 12866

This section addresses comments that EPA failed to comply with the provisions of E.O. 12866.

(1) *Comment*: Commenters assert that EPA erred by not complying with the requirement of E.O. 12866 to select among regulatory alternatives that are most cost-effective and maximize net benefits. Further, they say EPA did not examine alternative means to achieving its objectives that are more cost-effective, as it did in the SO_2 NAAQS decision (see API, IV-D-2247; State of N.C., IV-D-7003; NMA, IV-D-2158).

Response: For reasons discussed in section IV.A of the preamble to the final rule, the cited requirement of E.O. 12866 is inapplicable to NAAQS decisions. Moreover, the SO_2 NAAQS decision is not analogous to this rulemaking. In SO_2, EPA determined, based on its assessment of relevant scientific and technical information, that revisions to the SO_2 NAAQS were not appropriate for the reasons discussed in the preamble to the final rule (61 FR 25566; May 22, 1996). As in this case, EPA did not consider cost-effectiveness or the results of the Regulatory Impact Analysis in reaching its decision on the SO_2 NAAQS.

2. Regulatory Flexibility Act

This section addresses comments that EPA's failure to prepare a regulatory flexibility analysis and to convene a Small Business Advocacy Review Panel violates the Regulatory Flexibility Act as amended by the Small Business Regulatory Enforcement Act.

(1) *Comment*: A large number of commenters maintained that EPA's certification that the proposed revision to the PM NAAQS would not have significant economic impact on a substantial number of small entities and EPA's failure to prepare a regulatory flexibility analysis or convene a Small Business Advocacy Review Panel clearly violated the intent and plain language of the law. To support this position, several commenters presented extensive legal analysis (see NAM, IV-D-2274; ATA, IV-D-2245; AAMA, IV-D-2243).

Response: See section VIII.B. and section IV.A of the preamble to the final rule.

(2) *Comment*: It was also maintained 1) that the NAAQS itself will have significant impact on small business and that small business will bear a disproportionate impact; 2) that EPA's position that it cannot prepare a regulatory flexibility analysis is baseless (citing the PM RIA); 3) that EPA's informal, ad hoc overtures to small business are inadequate to satisfy SBREFA; and 4) that had EPA complied with statutory requirements, alternatives with less burdensome impacts on small business would have been identified (see ATA, IV-D-2245; NAM, IV-D-2274; AAMA, IV-D-2243; API, IV-D-2247; UARG,

IV-D-2250).

Response: See section VIII.B. and section IV.A of the preamble to the final rule. See also summary and response to comments for the small business outreach meetings.

(3) *Comment*: A commenter argued that the only possible and appropriate time for EPA to comply with the RFA as amended by SBREFA is at the NAAQS revision stage since EPA acknowledges that it will not perform an RFA analysis at the SIP approval stage and if it were to do so, conducting 50 different RFAs would result in bureaucratic duplication and inefficiency. The commenter argued that EPA cannot "segment" its analysis in order to completely avoid RFA requirements (AAMA IV-D-2243).

Response: See sections VII.B and IV.A of the preamble to the final rules. As noted therein, the Clean Air Act requires EPA to set a NAAQS and calls on States to develop and submit SIPs within a specified period of time after EPA issues the standard. Any "segmentation" that occurs thus results from the structure and requirements of the Clean Air Act and not from any EPA action or design. More importantly, the purpose of the RFA is to motivate federal regulators to design federal regulations in a way that fits the scale of the entities that will be subject to those regulations. That purpose cannot be served in the case of the NAAQS, since NAAQS simply define a level of air quality to be achieved everywhere in the country primarily through State regulation. Further, the RFA does not require or authorize EPA to disapprove a State's implementation plan because of the State's choice of sources to regulate. Fundamentally, the congressionally-designed mixture of Federal and State responsibilities for achieving clean air makes the RFA inapplicable to either setting or implementing the NAAQS, except to the extent EPA promulgates federal regulations establishing control requirements that will apply to small entities (e.g., reformulated gasoline standards).

(4) *Comment:* A number of commenters argued that EPA's claim that it cannot perform an RFA analysis is baseless and cited a variety of figures from EPA's RIA suggesting economic disruption or differential impact on small businesses (NAM IV-D-2274; API IV-D-2233; UARG, IV-D-2253; NMA, IV-D-2247).

Response: See sections VII.B and IV.A of the preamble to the final rules. As explained therein, EPA has attempted in the RIA to provide some insight into the potential impact on small entities of NAAQS implementation. In light of States' role in implementing the NAAQS, however, the RIA can assess only hypothetical State control strategies. As such, the RIA cannot and does not take the place of an RFA analysis, which is supposed to identify the types of small entities that will be subject to the federal rule being promulgated and ways of tailoring the rule to the size of the small entities being regulated. The RIA's small entity analysis, by necessity, depends on hypothetical State control strategies that may not occur and that EPA is not in a position to control.

3. Unfunded Mandates Reform Act

This section addresses comments that EPA failed to comply with the requirements of the Unfunded Mandates Reform Act (UMRA).

(1) *Comment*: EPA erred because it failed to comply with the requirements of UMRA. EPA is obligated to prepare a section 202 written statement, to conduct outreach efforts with small governments pursuant to a small government plan under section 203, and to solicit and evaluate input from State, local, and tribal officials under section 204. Finally, EPA's present failure to comply with UMRA is inconsistent with the SO_2 and NO_2 NAAQS decisions in which EPA did not disclaim application of UMRA. (NAM, IV-D-2274; ATA, IV-D-2245; AAMA, IV-D-2243; API, IV-D-2247).

Response: See section VIII.D and section IV.A of the preamble to the final rule. See also summary and response to key issues raised at outreach meetings with State and local officials.

IV. MISPLACED COMMENTS

A. Comments on implementation-related issues

1. Attainability of standards

This section addresses comments pertaining to the attainability of the proposed PM standards.

(1) *Comment*: Attaining a $PM_{2.5}$ standard would impose significant economic burdens, and such a standard may not be attainable in some areas (API, IV-D-2242).

Response: As discussed in section IV.A of the preamble to the final rule, the costs and technological feasibility of attaining ambient standards are not to be considered in setting them.

2. $PM_{2.5}$ Monitoring Issues

This section addresses comments that the $PM_{2.5}$ monitoring program will have a financial impact on the States.

Comment: Several commenters expressed concern about the financial burden that establishment of a $PM_{2.5}$ monitoring program would impose upon the States (IES Industries, IV-D-2150; Maricopa Co., Arizona, IV-D-2227).

Response: The preamble to the final 40 CFR part 58 requirements addresses comments

concerning the cost of the $PM_{2.5}$ monitoring network.

2. Implementation issues

A number of commenters submitted comments regarding implementation issues that are not relevant to the PM NAAQS review. Therefore, they are not being responded to in this document.

B. Comments on Regulatory Impact Analyses

This section addresses comments concerning the adequacy of the Regulatory Impact Analysis (RIA).

(1) *Comment*: The RIA is inadequate because it assesses the cost of only partial attainment of the proposed standards. In addition, the benefit estimates were artificially high. Further, the analysis is incomplete because it fails to analyze the full range of control measures likely to be imposed on the transport industry, does not assess indirect impacts (e.g., increased fuel costs), and does not assess the cost of administrative burdens (ATA, IV-D-2245; AAMA, IV-D-2243).

Response: Because the costs of implementation cannot be considered in setting or revising ambient air quality standards (see section IV.A of the preamble to the final rule), the RIA was not considered in EPA's decision on the standards. For the same reason, comments on the RIA were not considered in the decision. Comments on the draft RIA were considered, as appropriate, in developing the final RIA.

APPENDICES

Expanded Responses to Scientific and Technical Comments
on the Health Effects Evidence for Particulate Matter

U.S. Environmental Protection Agency

Research Triangle Park, N.C. 27711

July, 1997

EXPANDED RESPONSES TO SCIENTIFIC AND TECHNICAL COMMENTS ON THE HEALTH EFFECTS EVIDENCE FOR PARTICULATE MATTER

The purpose of this set of appendices is to provide expanded discussion and analyses in response to the large volume of public comments received on scientific and technical issues associated with the use of available scientific information on the health effects of particulate matter (PM) in revising the National Ambient Air Quality Standards (NAAQS) for PM. The Appendices were prepared by EPA staff in the National Center for Environmental Assessment (NCEA) and the Office of Air Quality Planning and Standards (OAQPS) in Research Triangle Park, NC. These two offices are responsible for preparing the PM Criteria Document (CD) and Staff Paper. Most of the comments received on scientific and technical issues addressed various aspects of community epidemiological studies of the health effects of PM. Appendix A provides responses to comments on general issues regarding the use of epidemiological studies in standard setting. In Appendix B, additional responses are provided to key issues raised by commenters regarding specific studies available as part of the criteria review. Newer studies published after completion of the criteria review are discussed in Appendix C. Appendix D addresses comments on the important issues of exposure misclassification, the role of indoor air pollution, the relationship of ambient air pollution to personal exposure, and measurement error in community studies.

APPENDIX A. RESPONSES TO COMMENTS ON GENERAL ISSUES REGARDING THE USE OF EPIDEMIOLOGICAL STUDIES IN STANDARD SETTING

I. INTRODUCTION

One body of commenters provided comments on EPA's use of PM epidemiological studies as a general matter, mainly citing criteria set forth by Hill (1995) or other authoritative sources as a means to judge the strength of epidemiologic evidence in supporting conclusions regarding possible cause-effect relationships between ambient PM exposure and reported mortality and morbidity effects. Most of these commenters appeared to approach the issue from the perspective of the use of epidemiology for making scientific and policy conclusions in occupational settings or in the kind of hazard assessment framework used in addressing carcinogens. This perspective is somewhat different from the case of PM and other criteria air pollutants, where community epidemiology is often central to the scientific and policy assessment. Assessing carcinogenic risk often involves extrapolation of results from high-dose animal experiments to humans, and occupational studies also often have short-term exposures at higher concentrations than the ambient exposures measured in community studies. Provided

below are responses to key points or issues raised in these comments, starting first with responses to the Gamble and Lewis (1996) published evaluation, which also commented specifically on more than 20 selected short-term PM epidemiology studies. This paper is addressed first, both because it is the most detailed discussion of these issues, and because it was cited and submitted by many commenters in support of their positions. The Gamble and Lewis discussion is followed by responses to general comments raised by one or more other commenters (e.g., the EOP Group on behalf of API, ENVIRON on behalf of Kennecott, the Sapphire Group, Cox Associates, etc.), some of which overlap issues raised by Gamble and Lewis (1996). The format of this Appendix, as well as those of Appendix B and Section I of Appendix C, follows the organization of the paper or comment under discussion. This leads to some duplication of issues and responses, which in many cases has been addressed by cross referencing to earlier responses.

II. RESPONSE TO COMMENTS

A. Review of Gamble, J.F and Lewis, R.J. (1996) Health and Respirable Particulate Matter (PM_{10}) Air Pollution: a Causal or Statistical Association? *ENVIRON. HEALTH. PERSPECT.* 104 (8): 838-850

Specific comments or points made in the Gamble and Lewis (1996) paper (G&L) are first identified below (by page #, column #, and line # in their 1996 publication), followed by the EPA response. References noted in G& L comments are cited by their number in the reference list of the published G&L 1996 paper.

1. G&L P 838, C 2, L 8-11: "Because correlation does not prove causation in observational studies, it is necessary to evaluate these associations using Hill's criteria (1)."

RESPONSE: Over the years, the type of criteria cited by Hill (1965) and others have provided a useful framework for organizing thinking about what epidemiologic observations of associations between variables may imply with regard to underlying relationships between the variables (e.g. chance or coincidental covariation or cause-effect). As such, the criteria cited by Hill are useful in helping to gauge the strength of evidence that exists at a given point in time, upon which judgments are made regarding the likelihood that epidemiologically-demonstrated associations reflect actual cause - effect relationships -- judgments that often serve as the basis for making informed decisions on appropriate medical treatment for individuals or for taking preventive measures to protect the public health of human populations.

Obviously, the more of the criteria cited by Hill that are met and the better that each is met, the stronger the confidence in the judgment that reported epidemiologic associations reflect cause-effect relationships and that medical or public health protective steps based on them are likely to have beneficial outcomes in addressing the health problems implied by the epidemiologic data.

However, as noted in Ch. 12 of the PM CD, neither Hill's criteria nor other lists of criteria should be used as an absolute checklist for establishing a cause-effect relationship, since any given criterion must be judged in context. As Hill himself stated on p. 299, "What I do not believe -- and this has been suggested -- is that we can usefully lay down some hard-and-fast rules of evidence that *must* be obeyed before we accept cause and effect. None of my nine viewpoints can bring indisputable evidence for or against the cause-and-effect hypothesis and none can be required as a *sin qua non*." [Bradford Hill's italics].

In fact, each criterion is discussed at length by Hill (1965) and exceptions or qualifications given for each of them. For example, some informative excerpts from Hill (1965) include the following:

• On Causation

"I have no wish...to embark upon a philosophical discussion of the meaning of 'causation.' ...with the aims of occupational, and almost synonymously preventive, medicine in mind the decisive question is whether the frequency of the undesirable event B will be influenced by the change in the environmental feature A. *How* such a change exerts that influence may call for a great deal of research. However, before deducing 'causation' and taking action we shall not invariably have to sit around awaiting the results of that research. ..."

Disregarding then any such problem in semantics we have this situation. Our observations reveal an association between two variables, perfectly clear-cut and beyond what we would care to attribute to the play of chance. What aspects of that association should we especially consider before deciding that the most likely interpretation of it is causation?"

• On Strength of Association

"First upon my list I would put the strength of the association. ... In thus putting emphasis upon the strength of an association we must, nevertheless, look at the obverse of the coin. We must not be too ready to dismiss a cause-and-effect hypothesis merely on the grounds that the observed association appears to be slight. ... Relatively few persons harboring the meningococcus fall sick of ... meningitis. Relatively few persons ... exposed to rat's urine contract Weil's disease. ... Nevertheless whether chance is the explanation or whether a true hazard has been revealed may sometimes be answered only by a repetition of the circumstances and the observations. ... [On smoking] The lesson here is that broadly the same answer has been reached in quite a wide variety of situations and techniques. In other words we can justifiably infer that the association is not due to some constant error or fallacy that permeates every inquiry."

• On Consistency

[Quoting Richard Doll on lung cancer in nickel refiners] "No causal agent of these neoplasms has been identified. Until recently, no animal experimentation had given any clue or any support to this wholly statistical evidence. Yet I wonder if any of us would hesitate to accept it as proof of a grave industrial hazard?"

• On Specificity

"In short, if specificity exists, we may be able to draw conclusions without hesitation; if it is not apparent, we are not thereby necessarily left sitting irresolutely on the fence."

• On Plausibility

"It will be helpful if the causation we suspect is biologically plausible. But this is a feature I am convinced we cannot demand. What is biologically plausible depends upon the biological knowledge of the day."

• On Coherence

"Nevertheless, while such laboratory evidence can enormously strengthen the hypothesis and, indeed, may determine the actual causative agent, the lack of such evidence cannot nullify the epidemiological observations in man. Arsenic can undoubtedly cause cancer of the skin in man but it has never been possible to demonstrate such an effect on any other animal. ... [On John Snow's epidemiologic investigation of cholera] Yet the fact that Koch's work was to be awaited another thirty years did not really weaken the epidemiological case though it made it more difficult to establish against the criticisms of the day - both just and unjust."

• On Experiment

"For example, because of an observed association some preventive action is taken. Does it in fact prevent? ... Here the strongest support for the causation hypothesis may be revealed."

2. G&L P 838, C 2, L 21-22: "... burden of proof regarding a causal association ..."

RESPONSE: The causal relationship of adverse human health effects due to exposure to ambient air containing high concentration of airborne particles was established beyond any reasonable scientific doubt by high-level incidents such as those observed in the Meuse Valley of Belgium in 1930, in Donora, PA in 1948, and in London in 1952. As one example, for the week of the well-documented 1952 London air pollution episode, the relative risk for mortality for all causes was 2.6, whereas that for bronchitis was as high as 9.3 and hospital admissions increased by a factor of two (Ministry of Health, 1954). In fact, these and other later severe air pollution episodes in London and other large cities tend to point toward a likely causative role of fine

particles, given that larger, coarse-mode particles settle out of the air and fine particles predominate the high concentrations of ambient PM that occur under low wind speed weather conditions that help to cause episodes. We are no longer in a situation where medical evidence of illness and death from PM exposure is so obvious as to admit no other explanation. The existence of a quantifiable concentration-response relationship between British Black Smoke (BS) and excess mortality was established in the 1980's. In the 1986 PM/SO_x CD and Schwartz and Marcus (1990), the PM effect was shown to exist even after adjustments for winter weather and for SO_2. Schwartz and Marcus found no evidence for a threshold in the relationship between PM and excess mortality extending down to concentrations lower than 150 $\mu g/m^3$ BS. The real problem, then and now, is detection and quantification of the relevant concentration-response relationship at much lower concentrations of PM than currently exist in most developed countries.

3. G&L P 838, C 2, L 37-38 and C 3, L 1-7: "The interpretation for a causal association between acute health endpoints and PM_{10} is based on correlation studies, which in epidemiology are called ecological studies because no measures of personal exposure are available (only group exposure data). For acute mortality and morbidity, time-series studies using a 24-hr sampling period for PM is the relevant type of study"

RESPONSE: The characterization of time-series studies here and at many other places in the G&L paper (see below) as *ecological studies* obscures a key strength of the PM time-series analyses. While it is true that no measures of individual exposure to PM_{10} or FP are available for these studies, EPA believes that variations in average population exposure to ambient PM in a given study location can be adequately represented by one or several community monitors of ambient PM and that the variations in average population exposure is the most appropriate PM measure to assess against daily variations in the city- or SMSA-wide mortality or morbidity data typically used in the time-series analyses. See also further discussion in Appendix D and Chapter 7 of the PM CD.

Daily time series studies of mortality, hospital admissions, and other health effects involve observations of effects in populations rather than in individuals. The predictors of effects include observations of air pollution and weather at one or a few locations in the SMSA, rather than for each individual. In this sense -- and only in this sense -- these studies may be described as "ecologic." The specific character of these time series analyses allows for much greater control of potential confounding factors than is the case in many other population-based study designs, e.g. cross-sectional ecological studies across different SMSAs.

The main difference is that the time series studies are based on changes in exposure across relatively short time segments, and the effects of exposure are compared within a single SMSA, rather than across different SMSAs. Many of the recent time series analyses reviewed in the PM CD and the Staff Papers are based on generalized linear models in which the number of events (people with adverse health responses) in one day is assumed to have a Poisson or hyper-Poisson

distribution. Control for extraneous cycles in the data unrelated to air pollution and for potential covariate effects is typically accomplished in the following ways: (i) long-term cycles are adjusted by removing time trends and seasonal variations, using a variety of strategies including non-parametric smoothing, parametric models for time trends, and seasonal adjustments by Fourier series, dummy variables, or seasonal stratification; (ii) other environmental variables that change on a time scale similar to the time scale for air pollution changes, such as weather variables, are adjusted by non-parametric smoothing, parametric models, or dummy variables; (iii) adjustments for many personal characteristics commonly found important in epidemiology studies (including age, race, gender, location of residence, location where death occurred, etc.) are made based on commonly used NCHS death certificate data. Most studies use age as a stratifying factor, but a few recent studies have involved other variables.

Episodes with high PM concentrations are likely to be associated with weather conditions that may also result in high concentrations of other pollutants. However, since seasonal and long-term changes in weather and emissions sources are likely to occur over the several years of a daily time series study, there are situations in which health effects of weather and at least some other pollutants can be separated. For example, ozone and high temperature effects in summer may be separated from winter weather effects with lower ozone concentrations, as shown by the Philadelphia analyses by Samet et al., (1996a,b)

The main advantage of daily time series studies relative to other so-called ecologic study designs is that it is possible to examine *differences of health effect in response to differences in air pollution exposure in essentially the same population over short periods of time.* During a period of a few weeks or possibly even a few months, the temporal changes in air pollution can be correlated with temporal changes in health effects in nearly the same population (or population sub-group) and over a generally similar range of weather conditions. These effects may then be statistically aggregated over many years, which also helps to avoid confounding with short-term events such as flu epidemics that may occur during some episodes, but not all episodes. Daily time series study designs therefore allow for a much higher degree of adjustment for potential confounders than do other population-based study designs.

4. G&L P 839, C 2, L 5 to P 840, C 1, L 36: "**Temporality**. Does cause precede effect? Are the time relationships plausible?"

RESPONSE: The information presented by Gamble and Lewis in fact tends to support the argument that daily time series studies may readily detect *changes* in health effects in a community in response to *changes* in air pollution during the few days (1 up to 3, 4, or 5) preceding the day on which the response occurs. In fact, by omitting responses with longer lags, daily time series studies may actually underestimate the effects of PM exposure. As discussed below, the analyses of temporality invariably favor changes in air pollution exposure preceding changes in health effects indicators, which clearly meets the causal hypothesis temporality

criterion. The range of observed time lag is also generally plausible, but may differ from cause to cause and from place to place across different studies.

5. G&L P 839, C 3, L 4-9: "When PM concentrations for the same day are used, it is not clear that there is enough time for exposure-related deaths to occur, especially if the deaths occurred before the full day's exposure is completed."

RESPONSE: Immediate same-day mortality or severe morbidity (e.g. respiratory distress or heart attacks) cannot be entirely ruled out as asserted here by G&L. During episodes of high air pollution, some cardiorespiratory deaths attributed by medical observers to air pollution were known to occur outside of the home ("on the street") and, at times, rather quickly. In particular, death from cardiovascular disease is likely to occur suddenly. Whether and how low concentrations of ambient PM might cause rapid deterioration in respiratory or cardiovascular health (e.g. via possible cascading effects of particle-induced lung inflammation or, perhaps, cardiac arrhythmia, hinted at by recently reported animal toxicology findings) remains to be more definitively demonstrated. The more germane point to focus on here is the difficulty in knowing, with great certainty, that effective PM exposure on a given day actually preceded the health effect on the same day, arguing for great caution in accepting same-day lag (zero-lag) analyses as providing meaningful information. Results from analyses using longer (e.g. 1 to 5 day lag times) are likely more credible.

6. G&L P 839, C 2, L22 and C 3, L 1-3, and 9-15: "These short lag times suggest PM may be lethal for persons already near death and who would have soon died even without increased PM exposure ... Except for persons already near death, it is not obvious that low-level PM_{10} concentrations could cause such quick and severe effects. While deathbed effects may have an appropriate time frame, this type of death cannot explain all the deaths attributed to PM (29)."

RESPONSE: While, as discussed below, it is possible that some PM-related deaths are "deathbed effects", there is as yet little concrete evidence on this point. It is also possible that individuals who are not bed-ridden or in hospital may be outdoors during PM episodes and die outside home or hospital, especially from cardiovascular causes. Indeed, hospitalized individuals receiving treatment for respiratory or cardiovascular diseases during a pollution episode may be less likely to succumb than those not in hospital. The temporal progression of death or illness may not be as simple as the speculations in the G&L paper suggest. For example, some individuals whose health status is compromised during an air pollution episode may indeed die or may be admitted to hospital long after the episode is over, during a period of lower air pollution, and so are counted as part of the "baseline" for relative risk evaluation.

The degree of lifespan shortening associated with PM exposure in these studies is viewed by many as an important consideration in evaluating mortality effects in a public health context. The epidemiological findings of associations between short- and long-term ambient PM concentrations and premature mortality provide some insight into this issue. The mortality

effects estimates associated with long-term PM concentrations in the prospective-cohort studies are considerably larger (Six City study) to somewhat larger (ACS study) than those from the daily mortality studies, suggesting that a substantial portion of the deaths associated with long-term PM exposure may be independent of the deaths associated with short-term exposure (U.S. EPA, 1996a, p. 13-44). The Criteria Document suggests that the extent of lifespan shortening implied by the long-term exposure studies could be on the order of years (U.S. EPA, 1996a, p. 13-45).

As discussed in the Staff Paper, attempts to quantitatively evaluate the extent of lifespan shortening in the daily mortality studies to date provide no more than suggestive results, with the investigators recognizing that more research is needed in this area (U.S. EPA, 1996b, p. V-19-20). The limited analyses available suggest that at least some portion of the daily mortality associated with PM may occur in individuals who would have died within days in the absence of PM exposure (U.S. EPA, 1996b, p. V-19-20). Researchers in this area also note that it is possible that the reported deaths might be substantially premature if a person becomes seriously ill but would have otherwise recovered without the extra stress of PM exposure (U.S. EPA, 1996b, p. V-19-20).

Quantification of the degree of lifespan shortening inherent in the long- and short-term exposure mortality studies is difficult and requires assumptions about life expectancies given other risk factors besides PM exposure, including the ages at which PM-attributable deaths occur and the general levels of medical care available to sensitive subpopulations in an area. Because of these uncertainties, it is not possible to develop with confidence quantitative estimates of the extent of life-shortening accompanying the increased mortality rates that have been associated with exposures to PM (U.S. EPA, 1996a, p. 13-45).

7. G&L P 839, C 3, L 16-23: "Increased susceptibility to infectious diseases ... is a possible cause of PM-increased mortality among elderly people who have cardiorespiratory disease. (30). Because these diseases develop and evolve over days and weeks, a 24-120 hr time frame is too short to meet this criterion (29)"

RESPONSE: As noted above, individuals who die or are admitted to hospital long after being exposed to high PM concentrations are likely to be counted among the "baseline" levels of mortality and morbidity, against which deaths during a PM episode are compared. Therefore, counting excess deaths or illnesses only 1 to 5 days after PM exposure may underestimate the total PM effect. Very few studies have examined the question of delayed PM effects. As noted in the Staff Paper (p. V-22), three such studies reported statistically significant associations between community air pollution, as indicated by PM, and work loss days and restricted activity days (Ostro, 1983; Ostro and Rothschild, 1989; Ostro, 1987). More specifically, the study by Ostro and Rothschild (1989) reported significant associations between PM exposure and respiratory-related restricted activity days. All of these studies used two- to four- week lag times between elevations in PM levels and school absences, work loss days, and restricted activity

days. This suggests that not only are there immediate effects after elevations of PM exposure (e.g., increased hospital admissions), but PM may elicit effects which are exhibited at a later time. These results of these studies are consistent with a hypothesis of increased susceptibility to respiratory infection resulting from exposure to PM.

Moreover, even very short timeframes (24-120 hr) cannot be ruled out a possibly applicable for exacerbation by ambient PM components of overall constellations of pathophysiological problems (e..g, conditions associated with existing respiratory infection) that might rapidly deteriorate, for example, due to PM-induced lung inflammation or cardiac arrhythmia.

8. G&L P 839, Table 1, Study 5, Birmingham AL, 1985-1988 (6): "PM_{10} (0-4 days)"

RESPONSE: The analyses cited in the PM CD used the best-fitting time span, with an average of the PM_{10} concentrations observed 1, 2, and 3 days prior to mortality. Other analysts have noted that current-day PM_{10} has little (or even negative) relationship to mortality in Birmingham.

9. G&L P 840, C 1, L 1-16: "A third major possible cause of death due to PM is exacerbation of underlying cardiac or pulmonary disease. ... Although many susceptible patients with cardio-respiratory disease smoke until late in their disease, smoking does not result in acute hospital admission (29)."

RESPONSE: There is no concrete evidence presented by G&L to support this speculative contention (that smoking does not result in acute hospital admission) and the paper by Gamble and Lewis does not present any evidence demonstrating that their argument is relevant to the analysis and interpretation of the PM daily time series studies.

10. G&L P 840, C 1, L 17-22: "Cardiovascular disease (CVD) deaths may be misclassified as respiratory deaths (29). When such misclassification occurs, the time course for respiratory disease death does not appear to be appropriate to the 0-5 day lags of time series studies."

RESPONSE: It is more likely that the respiratory deaths would be misclassified as CVD deaths than vice versa. Also, the G&L paper presents no evidence that misclassification of disease category is correlated with air pollution levels. Therefore, misclassification cannot be a confounder of the health effect estimates in time series studies, which are based on *changes* in response related to *changes* in air pollution after adjustment for other time-varying factors.

11. G&L P 840, C 1, L 23-36: "The time period also may not be appropriate for morbidity-related effects. ... Canny et al. (30) reported ... symptoms about 41 hr before arrival ... The duration of symptoms was longer than 72 hr for 16% of the patients. Major precipitating causes included respiratory infection (75%) ..."

RESPONSE: The statements are in fact consistent with a detectable *change* in hospital admissions associated with *changes* in air pollution in the preceding 1 to 4 days, and with less

detectable responses over a longer period of time that are added to the baseline or non-episode incidence. The association with respiratory causes in pediatric asthma patients is also not inconsistent with PM as a possible contributing factor.

12. G&L P 840, C 2, L 1-4, and C 3, L 9-14; P 841, C 1, L 1-4: "**Consistency** ... separate analyses of the same populations by different investigators have produced inconsistent results that are contrary to those of the original authors ... results are dependent on the model used."

RESPONSE: The paper by Gamble and Lewis apparently accepts all model-building strategies by all authors as equally valid and worthy of consideration. According the same degree of credibility to all studies, without regard to their scientific merit and in some cases to those suffering serious methodological problems, is not particularly helpful in resolving complex scientific and regulatory issues. EPA evaluated all of the studies that were available and found some to be more reliable than others. Section 12.6 of the PM CD, as reviewed and closed on by CASAC, provides a much more detailed assessment of the models, with particular attention to directed model-building strategies that are specified in advance of the analyses so as to minimize the risk of chance findings that are inconsistent with prior scientific knowledge. The models developed by the Health Effects Institute (HEI) for Philadelphia in their Phase IB studies (Samet et al., 1996a,b) also illustrate well some generally adequate model-building strategies.

EPA assessments of model results examined not only the model finally selected by the investigator, but also the strategy used to select the model, usually including initial exploratory models aimed at characterizing underlying weather-related and/or other important covariate contributions to the overall structural relationships. These exploratory approaches acknowledge that there does not yet exist a fundamental theoretical or experiential basis for *a priori* definition of any single correct bioclimatological or biometeorological model. There may indeed be real differences among regions in terms of demographics that affect health responses to weather. The demographic basis for health effects of weather may change over time, and acclimatization may also account for regional differences in response to similar weather conditions. But, there is now enough evidence showing that there are differences in response to weather conditions, and that the overall response (seasonally adjusted or not) is nonlinear, with increasing adverse health effects for very low temperatures and for very high temperatures. Also, lag structures for weather effects may well vary with season: for example, a 1 day lag in summer and 3 day lags in winter. Studies that did not allow sufficient flexibility to capture such differences should be given less credibility than those studies that allow enough flexibility to fit a good baseline model.

The second stage in appropriate modeling is to systematically evaluate air pollution effects. A directed search procedure is also desirable here. Directed searches should start from prior scientific knowledge about pollutant effects. The London mortality studies suggest that there may be a near-linear relationship extending to lower concentrations than the current NAAQS, and this relationship is more likely to be detected using thoracic or respirable particle indicators such as PM_{10}/PM_{15}, or indicators for smaller particles such as $PM_{2.5}$ and sulfates, since the London BS index was believed to correspond roughly to a D_{50} cutpoint of $PM_{4.5}$ with a few

larger particles (up to 7-9μm) being sampled. Other pollutants may also be evaluated, taking into account various factors that are responsible for more or less correlation among concentrations of various air pollutants.

Directed searches with PM and other air pollutants should also include only plausible models for pollutant effects. The PM lag structures suggested by previous studies generally include the prior day and second previous day concentrations. Some studies find greater predictiveness when longer lags are also included, 3 or 4 days prior to the health effect. Other studies find greater predictiveness when the current day is included as well as 1 or 2 prior days. In general, there is less evidence for current-day exposure and for long-ago exposures than for some combination of concentrations with 1-day to 5-day lags. Most investigators have used simple moving averages, and only a few studies have considered weighted moving averages of prior exposure. (This should not be confused with the practice of subtracting 15-day or 19-day moving averages that include future exposures as well as current and past exposures). There is a possibility that different lag structures may in fact better characterize different SMSAs, reflecting differences in statistical power for lag structure detection in different studies with different ranges of PM levels, different weather conditions, and different co-pollutants. There may also be some relationship to real differences in the duration of synoptic categories in different SMSA's or to differences in the duration of air pollution episodes among SMSAs.

An optional third stage in model building would reexamine the significance of effects in the baseline model after air pollution has been added, and would delete the most non-significant effects after air pollution has been included in the model. Comparison of non-hierarchical final models with the baseline model(s) should not be difficult.

The least desirable model-building strategy is one in which a factorial design is laid out *a priori* and then all models within the factorial design are fitted to the data. This strategy allows little or no ability to consider the implausibility of many combinations or permutations included in the purely hypothetical factorial pattern; and rote adherence to simply running every combination, no matter how implausible, precludes learning from the data and production of better models based on the data that may not have been included in the original design.

Specific issues in the approaches used by different investigators are discussed below. These constitute critical elements in the use and overall interpretation of divergent findings from different investigators, and appear to have been ignored in the Gamble and Lewis paper. Differences of interpretation of the various analyses constitute a normal element of scientific discourse, and do not indicate per se that the results are "incoherent" especially when some analyses are more methodologically credible than others.

13. G&L P 841, C 1, L 5-20: "**Steubenville, Ohio** ... Moolgavkar et al. (34) ... found that TSP was not significant when SO_2 was included in the regression model. ... the results were not robust ..."

RESPONSE: EPA discussed the Stuebenville TSP analyses, including the original analyses by Schwartz and Dockery (1992b) and the subsequent reanalyses by Moolgavkar, in detail in Sec. 12.6 of the PM CD, but did not place great reliance on any of them because TSP was a much less adequate indicator of thoracic particle exposure than PM_{10}. Furthermore, there was a possibility of high correlation between TSP and SO_2 because of the major role played by stationary combustion sources in Steubenville during the years of these studies. A more recent analysis of Steubenville data (Schwartz, Dockery, and Neas, 1996) suggests significant health effects from $PM_{2.5}$ (FP). As discussed in the staff paper (EPA, 1996b), in situations where it is difficult to separate out the effects of PM and SO_2, there are several reasons (related to indoor penetration and respiratory tract penetration) to conclude that PM is likely the active agent. Moreover, similar PM associations occur in multiple areas where SO_2 concentrations are high or low (EPA, 1996b; Figure V-b Staff Paper).

14. G&L P 841, C 1, L 21 to C 3, L 14. "**Philadelphia, Pennsylvania** ... Li and Roth (10) added 10 more years of data ... TSP was not significantly associated with any cause-specific mortality ... Moolgavkar et al. (35) concluded that, because the copollutants were so highly correlated, it was not possible to single out any specific pollutant effect. Wyzga and Lipfert (18) reported on 18 years of data ... Li and Roth (7) reported mixed results ... "

RESPONSE: EPA discussed these analyses in detail in Sec. 12.6 of the PM CD. There was a possibility of high correlation between TSP and SO_2 because of the major role played by stationary combustion sources in Philadelphia during the years of these studies. Moreover, as noted by several CASAC panelists with expertise in epidemiology (Lippmann et al., 1996), these studies are largely superseded by a more recent analysis of Philadelphia TSP data (Samet et al., 1996a, b). These analyses, examined at length in the PM CD suggest that significant health effects are associated with TSP and SO_2, with ozone, and with lagged CO, but that TSP/SO_2 effects can be reasonably separated from effects of the other pollutants. Robustness of these findings was evaluated in a number of different ways, and the effects of air pollution were not substantially confounded by methods used for long-term trend or weather adjustment. The findings by Samet et al. are regarded by EPA and the CASAC epidemiological experts as being more definitive than those relied on in the Gamble and Lewis paper.

Gamble and Lewis describe the various Philadelphia analyses as "a similar situation of conflicting results were observed from five studies conducted in Philadelphia." Only two of the studies are shown in their Table 1 with study details, but the others are referenced but not evaluated. This differs from the extensive and more recent review of the Philadelphia studies in the CD, particularly the extended analysis of the most recent work by Samet et al. (1996a,b). It also differs with the most recent HEI Oversight Committee comments on these studies (HEI, 1997), which state that:

> "Although individual air pollutants (TSP, SO_2, and ozone) are associated with increased daily mortality in these data, the limitations of the Philadelphia data make it impossible to establish that particulate air pollution alone is responsible for the widely observed

> associations between increased mortality and air pollution in that city. All we can conclude is that it appears to play a role" "[HEI, 1997; p.38.]

The oversight panel goes on to say that in this situation:

> "Consistent and repeated observations in locales with different air pollution profiles can provide the most convincing epidemiological evidence to support generalizing the findings from these models. This has been the approach reported by the EPA in its recent Criteria Document and Staff Paper."[HEI, 1997; p.38.]

Gamble and Lewis do not do that. They attempt to identify in their discussion differences between analysis of a data set by different groups and point at this as proof of inconsistency. They fail to integrate the entire epidemiology data set and make balanced inferences from it.

15. G&L P 841, C 2, L 11-13: [Li and Roth (10)] "For nearly every positive result, there is a negative or nonsignificant result pointing in the opposite direction."

RESPONSE: This statement reflects uncritical acceptance by Gamble and Lewis even of non-directed or misdirected model specification comparisons. Most of the model specifications run by Li and Roth were not evaluated in an appropriate specification search, and appropriate preferential weight was not given to models concordant with prior knowledge. For example, Li and Roth ran numerous model specifications using same day (zero) lags, which Gamble and Lewis themselves reject as likely inappropriate (see above). Several CASAC panelists with expertise in epidemiology (Lippmann et al., 1996) characterized the analysis of the Philadelphia data et by Li and Roth (1995) as purporting to show that a panoply of seemingly conflicting findings is produced with different modeling strategies; they further point out that this paper is superseded by the HEI report, which shows conclusively that the confounding effect of weather was appropriately controlled in the original analysis, and that the original results are not an artifact of the modeling strategy.

16. G&L P 841, C 3, L 12-14, from Samet (37): "Assessment of the causality of associations should not rest solely on model results."

RESPONSE: EPA's assessment does not so rest. Causality of the relationship between cardiorespiratory effects and airborne particle exposure was convincingly demonstrated by medical observation during several mid-century severe air pollution episodes. While EPA acknowledges that it is difficult to quantify the concentration-response relationships at PM levels below the current NAAQS, evidence presented in the 1986 EPA PM/SO_x CD establishes a quantifiable relationship based on BS in London, not wholly confounded by SO_2 or by London winter weather, that appears likely to extend below the current PM_{10} NAAQS. Causality at such low PM levels appears likely, based in part on the existence of a spectrum of effects, from excess mortality and frank cardiorespiratory illness associated with particle exposure at very high levels, to quantifiable concentration-response relationships in London winters for the years

1958-1973 even at much lower BS levels, to associations that have been detected in recent studies even when PM levels are below the current NAAQS. The Gamble and Lewis paper appears to ignore the medical evidence of the early episodes, and to minimize the findings of more recent London analyses as well as the results of an intervention in Utah Valley (see below).

17. G&L P 841, C 3, L 29-31: "**London** ... high correlations (0.79-0.96) between BS and SO_2 make it impossible to distinguish their separate effects."

RESPONSE: This is discussed in the EPA's 1986 PM/SO_x CD and in Schwartz and Marcus (1990). In some years and in some analyses, the SO_2 effect appears to be more significant than the BS effect. However, regarding each of the 13 winters separately as independent studies for which a combined analysis is possible, the BS effect is positive and statistically significant in most years, whereas the average SO_2 effect is nearly equal to zero, with almost the same number of positive and negative estimates. While the high correlation between BS and SO_2 is a complicating factor in the analysis of some datasets, there are studies in the U.S. in which high levels of PM are found with low levels of SO_2, allowing incisive estimation of a PM effect.

18. G&L P 841, C 3, L 32 to P 842, C 1, L 26: "Ito et al. (38) ... found ... all three pollutants [BS, SO_2, aerosol acidity] were significant. However, no particular pollutant effect could be determined because of pollutant collinearity and lack of quantitative information about measurement error ... Lippman and Ito (39) reanalyzed the London mortality data (1965-1972) ... The strongest correlations were with SO_2 and H^+ (not BS) ... a new pollutant (H^+) has been added to the list of potential confounders, and PM was the least important pollutant".

RESPONSE: The interrelationships among PM, SO_2, sulfates, and H^+ were discussed in detail in Sec. 12.5 of the 1996 PM CD. These are clearly closely interrelated pollutants, with SO_2 a precursor to sulfates, with a substantial part of aerosol acidity (H^+) being related to airborne sulfuric acid from sulfates, and with sulfates contributing a substantial part of fine PM, which is in part indexed by BS (assuming some correlation between sulfates and BS). Thus, arguments can be made for all of these (BS, SO_2, sulfates, and H^+) serving to some extent as indicators of the fine particle fraction of ambient PM.

Gamble and Lewis have not explored the importance of causal relationships in assessing the role of potential confounders in an epidemiology study. In fact, in some widely used textbooks, the above-noted pollutants might not even be considered as properly defined confounders. K.J. Rothman (*Modern Epidemiology,* 1986, p. 94) cautions that "A confounding variable must not be an intermediate step in the causal path between the exposure and the disease." Since the concentrations of these pollutants are closely related, it is not clear whether the causal pollutant in this chain *that is closest to the adverse health effects* is H^+, sulfates, BS itself, or some combinations of these pollutants or the entire fine particle mix indexed by one or more of them. EPA observes that Gamble and Lewis have failed to identify which of the stated pollutants they

think are confounding which effects of which others, a matter which largely obscures the real inferential issues.

As Gamble and Lewis note later in the paper (P 846, C 3, L 12-15), "... the pollutant with the lowest measurement error will have the spuriously highest regression coefficient ..." If the latter statement is true, and the measurement error among London pollutants was unknown, then there is little basis for interpreting the relative significance of the BS, SO_2, and H^+ in view of their high correlation. As noted earlier, other approaches suggest that effects of BS and SO_2 can be separated to some extent, in which case H^+ (which may be the least accurately measured of the three) could be regarded as biasing the estimated effects of BS and SO_2.

19. G&L P 842, C 1, L 27-64: "**Utah**. Three studies were conducted in Utah, with different results for each. ... The differences do not seem to be due to differences in exposure between study areas because PM levels in the two counties were similar (average of 47 $\mu g/m^3$ for Utah County versus a median of 48 $\mu g/m^3$ for Salt Lake County. ..."

RESPONSE: Only two of the three "Utah" studies cited by Gamble and Lewis were actually done in the same location, the Utah Valley studies by Pope et al. (5) and by Lyon et al. (4). The study by Styer et al. (3) was conducted in Salt Lake County, a separate SMSA from the Utah Valley (Orem-Provo) SMSA. With regard to the latter paper, four CASAC panelists with expertise in epidemiology noted that "among papers considered as not supporting the main conclusion of the EPA criteria document, that of Styer et al. (1995) fitted separate regressions to each month of the year and found significant particulate effects only in a few of the months. But such partitioning of data in small time segments is considered to be inappropriate because it results in a significant loss of statistical power and thus a loss of sensitivity to the moderate relative risk associated with ambient air pollution and a loss of ability to separate the effects of one pollutants as opposed to another (Lippmann et al, 1996)."

The data from Utah County have been analyzed by three separate research groups. Gamble and Lewis have overlooked or were unaware of the reanalysis of the data from Pope et al. (1992) by the Health Effects Institute (Samet et al., 1995). All of the available studies for the Utah Valley were extensively evaluated in Section 12.6 of the PM CD, in addition to (4) and (5). In fact, the results of both reanalyses, as well as subsequent analyses of these data by Pope and Kalkstein (1996), supported the original findings of Pope and colleagues. Pope et al. (1992) found significant associations with respiratory, cardiovascular, and total mortality; total mortality was found to increase 16% with a 100 $\mu g/m^3$ increase in PM_{10}. Increases of 43%, 20% and 5% were predicted for respiratory, cardiovascular and all-other-causes deaths per 100 $\mu g/m^3$ increase in PM_{10}. Under the Health Effects Institute's Particle Epidemiology Evaluation Project, Samet et al. (1995) reanalyzed these data using a carefully constructed modeling strategy to control for weather factors and other trends. The results of this analysis essentially replicated the original findings. Increases of 48%, 20%, and 17% were found for in respiratory, cardiovascular and total mortality with a 100 $\mu g/m^3$ increase in PM_{10}; all associations were statistically significant.

Gamble and Lewis also fail to note that Lyon et al. (1995) themselves corraborated the Pope et al. findings when using PM as a continuous variable. The authors state: "We confirmed the previous finding that exposure to fine particulate air pollution (particle diameter of $\leq$ 10 um) in amounts of 50+ $\mu g/m^3$ increased daily mortality by 4%." Otherwise, the studies by Lyon et al. (4) investigated many interesting issues, such as location and cause of death, and age group. Some of the analyses appeared to involve stratification by population subgroups for the purpose of analysis, which generally reduces the power of a study to detect differences. The PM_{10} effect was characterized by a dichotomous variable, whether the PM_{10} concentration for the day was above or below 50 $\mu g/m^3$. Dichotomizing a continuous effect generally reduces the power of a study, especially if the cutpoint is selected arbitrarily. Although this study evaluated many interesting hypotheses, it is limited by its crude assessment of PM_{10} effect, by stratification and other analytical approaches that reduced the power of the study, and by a vague description of the model specification search strategy.

There were a number of subsequent reanalyses for Utah Valley by Pope and his coworkers that increase confidence in the robustness of the reported results. These included alternative models for weather, especially a comparative study of synoptic weather categories, nonparametric smoothing, and parametric alternatives published by Pope and Kalkstein in 1996.

20. G&L P 842, C 2, L 18-35: "**Birmingham, Alabama**. ... Li and Roth (7) ... found virtually no association of PM_{10} and mortality except when maximum temperature was used to control for weather. When a more appropriate variable for temperature (deviation from threshold) was used, the association disappeared. Because the association with temperature is not linear (e.g. a U-shaped relationship) ... the results were not robust, but sensitive to lag times, models, and temperature variables used in the analyses."

RESPONSE: These findings were presented at the May, 1996 meeting of CASAC and were independently reviewed by CASAC and by EPA staff. The comments of some CASAC members are particularly illuminating. We quote directly from the transcript for May 17, 1996:

> Dr. Kinley Larntz:. "What is very important ... is to actually analyze the experiment he [Dr. Roth] did and to see what the effects are. Some of the negative effects come from zero lags where you expect nothing to have occurred, so it is a negative effect; it should be noise, but it is a negative effect. When you mix those zero lag negative effects with positive effects at other lags, you might get a distribution like he had. ... I think the implication from that presentation was that you get whatever you would like from just messing around. I think that is not correct. Those analyses are not all independent (Transcript for May 17, 1996, P 37 L 20 to P 38 L 11)."
>
> Dr. Jan Stolwijk: "What he [Dr. Roth] did was to do a number of analyses, most of which somebody who was actually doing a study trying to find out about things would not do. It would predictably give us non-response. If you take the experimental design space ... and that experimental design space includes a large number of things that are not

> expected to produce a result, you are going to dilute the results that you seek. If you select out of his [Dr. Roth's] matrix the things that other people have done, he comes to a different conclusion than when he takes his whole matrix. ... you are going to get a random effect that shows that there is no effect. He [Dr. Roth] did this, I think, on purpose in this case. Most epidemiologists, I think, have been trained to limit their observations to something that they can state or would have stated before they started and observe that and base their conclusions on it." . . . "What he did was to do a series of experiments essentially or a series of analyses which most people who are knowledgeable in the area would say would not produce any result, we would not expect to see any result, and he [Dr. Roth] did not see any result. As a result, results that you would have expected to have seen did not stick out any more as notable findings (Transcript for May 17, 1996, P 45 L 15 to P 46 L 16)."

EPA's detailed responses to Drs. Li and Roth's comments (see Appendix C) expand on these CASAC points. Some of the more salient concerns with the Roth analyses can be briefly noted here. First, the 9600 computer runs fitted to Birmingham data by Li and Roth did not appear to follow any model specification search method or optimization scheme commonly used in statistical modeling. Rather, Roth basically ran numerous model specifications covering an enormous number of different combinations of pollutants, weather variables, etc. no matter how unplausible. For example, a number of the Roth model runs included same-day (zero-lag) PM concentrations, which the original investigators found to not be significantly related to mortality or morbidity effects. Other recent analyses of Birmingham mortality data by the National Institute of Statistical Sciences, submitted as part of their public comments and discussed elsewhere (Appendix C), also show a significant PM_{10} effect when the same 3-day PM_{10} moving average with lags 1 to 3 days is used (but not when same-day, zero-lag PM levels are included), even with a substantially different model for weather effects. These analyses suggest that the results of the original Birmingham mortality analyses (6) have some robustness against alternate models for weather, but (as noted by Larntz) moving averages for PM should not include same-day PM_{10} concentration.

Another area of concern is specification of the functional form for temperature effects modeling. While overall relationships between excess mortality and temperature are U-shaped over the year, these effects are experienced differently by season. A winter temperature-mortality relationship is likely to be flat over a range of comfortable ambient temperatures, but increasing with decreasing temperatures below some low-temperature change point or "threshold". A summer temperature-mortality relationship is also likely to be flat over a range of comfortable ambient temperatures, but increasing with increasing temperatures above some high-temperature change point or "threshold". The winter low-temperature threshold is much lower than the summer high-temperature threshold. Furthermore, it is highly unlikely that the winter effect per degree below the low-temperature threshold is the same as the summer effect per degree above the high-temperature threshold. The V-shaped temperature effects model proposed by Li and Roth is among the least plausible of all of the models that have been proposed by various investigators. Even if the "ideal" temperature had also been an adjustable parameter that had

been estimated from the data, the difference between the Li and Roth V-shaped function and a U-shaped function of temperature would include a substantial overestimation of the effect of temperature at temperatures somewhat above and somewhat below the "ideal" temperature. This is likely to introduce spurious confounding between weather and air pollution. Other methods, such as using indicator variables for temperature ranges or using nonparametric smoothing techniques, can provide more satisfactory models for weather effects than this V-shaped parametric model.

21. G&L P 842, C 2, L 51 to C3, L 6: "**Are Results Consistent When Using Different Study Designs?** ... inherent biases in ecologic risk estimates ... risk must be independently checked using individual-level study designs having personal exposure measurements. ... experimental or chamber studies provide almost the only available data that meet this requirement ... experimental studies do not show the risk from PM_{10} exposure suggested by the time-series studies of mortality and hospital admissions."

RESPONSE: EPA also regards experimental or chamber studies, or other health effects studies for individuals with appropriate personal or residential monitoring, as potentially valuable. However, such studies may provide little information that is directly comparable to the time series epidemiology studies. For example, Gamble and Lewis discuss possible "deathbed effects" of PM exposure on P 839. One assumes that they are not seriously proposing that such highly compromised individuals be used in chamber studies and be exposed to particle concentrations indicated by epidemiologic studies as being potentially lethal for them. Likewise, one can hardly imagine using somewhat less sick elderly subjects with pre-existing conditions such as CVD or COPD in chamber studies in which frank symptoms leading to hospitalization or death are possible.

The authors' characterization of time series studies as "ecologic" is inadequate and somewhat misleading, as discussed earlier. While group estimates of ambient PM exposure are provided by community monitors, the time series method uses *changes* in community PM levels as an indicator of *changes* in population exposure to ambient PM. For pollutants with a reasonably uniform regional spatial distribution, such as sulfates and $PM_{2.5}$, and PM_{10} in most cases, this assumption is appropriate. Other pollutants may be less adequately characterized.

In all experimental studies, including personal monitoring studies and chamber studies, it is not possible to compare single-subject results with population results unless the sample of subjects is representative of the population for which inferences are being made. It remains to be proven that subjects recruited for such studies are similar in important ways to those in the general population who are susceptible to PM-related effects. Subjects recruited into such studies may have a greater awareness of the adverse health effects of air pollution, or are in general more health-conscious than are other (possibly more susceptible) subjects. Selection problems are also likely to occur in retrospective studies on health status (for example, using HMO databases), and in prospective studies unless a rigorous recruitment effort is made. Gamble and Lewis do not note these study design issues.

22. G&L P 842, C 3, L 54-61: "**Strength of Association and Exposure-Response (E-R). ... Against.** Bias due to ambient concentrations that do not accurately reflect exposure does not always reduce the magnitude of the RR when present ..."

RESPONSE: As discussed in Chapter 7 of the PM CD and Appendix D below, ambient concentrations of sulfates and fine particles ($PM_{2.5}$) measured at stationary air monitors are likely to provide a good basis for estimating daily changes in exposure to sulfates and to FP generated by outdoor sources for most of the population in a region. While coarse fraction (CF) of thoracic particles ($PM_{10-2.5}$) may show somewhat greater exposure measurement error than FP, the overall characterization of PM_{10} exposure from outdoor sources by stationary air monitors is likely to introduce only small biases in RR estimates.

Gamble and Lewis are correct in suggesting that measurement errors do not always reduce the magnitude of relative risks. In multiple pollutant analyses, measurement error or, more generally, exposure misclassification, could theoretically bias effects estimates of PM or co-pollutants in either direction, introducing further uncertainties in the estimated concentration-response relationships for all pollutants (U.S. EPA, 1996b, pp. V-39 to -43). Relevant insights on this issue in material appended to public comments (Ozkaynak and Spengler, 1996) have prompted an expanded statistical analysis of the conditions under which such errors could inflate the magnitude of the effects estimates or the significance of PM relative to gaseous pollutants, as has been suggested by Lipfert and Wyzga (1995). This analysis, which appears in Appendix D below, finds that the conditions under which measurement error could inflate the effects estimates or significance of PM relative to other pollutants are restricted to a limited set of statistical relationships. Gamble and Lewis do not cite studies that suggest such conditions are likely to occur with respect to the measurement of ambient PM in relation to those for gaseous co-pollutants commonly used in epidemiological studies. Therefore, it appears unlikely that measurement and exposure errors for PM and other pollutants have inflated the estimated effects of PM, even in multivariate analyses. More importantly, the available evidence on the consistency of the PM-effects relationships in multiple urban locations, with widely varying indoor/outdoor conditions and a variety of monitoring approaches, makes it less likely that the observed associations of PM with serious health effects at levels allowed under the current NAAQS are an artifact of errors in measurement of pollution or of exposure (U.S. EPA 1996b, pp. V-39 to -43).

Measurement error biases in air pollution effect estimates are interesting from a scientific point of view, but may be irrelevant when the relative risk models are used for prediction. A very comprehensive summary of modern approaches to statistical inference in regression models when measurement error is present in the predictors has been published by Carroll et al. (1995). They conclude:

> "If a predictor X is measured with error, and one wants to predict a response *based on the error-prone version W of X*, then ... it rarely makes sense to worry about measurement error. ... The one situation requiring that we model the [distribution of] measurement

> error occurs when we develop a prediction model using data from one population, but we wish to predict in another population. A naive prediction model that ignores measurement error may not be transportable." (Carroll, D, et al., 1995)

As noted in their monograph, the complete characterization of exposure measurement error includes the measurement error structure of important covariates or modifying factors. Weather is one such factor. Since weather effects are likely to vary significantly from place to place, predictive models for one SMSA that include weather-related factors may not be transportable to some other SMSA. However, to the extent that coefficients for PM_{10} and $PM_{2.5}$ appear to be more similar to each other for different SMSA's than do coefficients for weather variables, the estimated effects for PM factors may be more transportable and useful for prediction.

23. G&L P 843, C 1, L 13-18: "... when confidence intervals around percentiles are available, as they are in the Health Effects Institute reanalyses (32), the lower 95% confidence intervals are mostly below 1, and the E-R trend is not obvious (see Fig. 1)."

RESPONSE: Grouping continuous E-R information into categories by percentiles generally involves a loss of information. This method was used by some of the Original Investigators, but has now largely been supplanted, even in HEI's own reanalyses (see Samet et al. 1996a,b) by use of linear pollution models or by flexible nonparametric smoothing functions. Gamble and Lewis plotted quantile results from (32) without noting an extremely important difference in methodology. From Samet et al., (1995), p. 14: "Our approach to performing analyses by quantile differed from that used by the Original Investigators, and our findings should be compared with those from the original analyses with this distinction in mind. We based the quantile groupings on the absolute levels of the particulate measures; the Original Investigators based their groupings on the residual values, after removing the effects of season and calendar time." Arguments can be made for either approach, but were overlooked by Gamble and Lewis.

24. G&L P 843, C 1, L 19-35: "Another way to assess the strength of association is to examine R^2 values, which measure how much of the variability in the observed data (e.g., mortality) is explained by the statistical model. ... Few air pollution studies have reported R^2 values. Those that have are summarized in Table 4."

RESPONSE: Few recent studies have reported R^2 (only one in Table 4 after 1990) because R^2 was recognized as an inappropriate goodness-of-fit index with data such as mortality counts or hospital admissions counts. Generally, each investigator sets up a limited menu of choices that are plausible (to him or her) and then tests those choices by fitting the adjustable parameters for each model and comparing the observed counts to the predicted values by a global goodness-of-fit statistic (denoted as GOF). Model evaluation should be based on an appropriate GOF.

Most investigators acknowledge that daily counts of deaths or hospital admissions should be modeled as Poisson or hyper-Poisson variables, with the Poisson deviance as an appropriate GOF. Log-normal models are still sometimes used because of the availability of standard off-

the-shelf software for Gaussian time series models, and may be a reasonable variance-stabilizing approximation for hyper-Poisson daily count data when the mean number of events is large (as in London). However, most studies that have looked at the variance have found that the variance inflation (residual variance/mean) is usually not much larger than the Poisson value of 1, so that log-transformed counts may not satisfy the homoscedasticity assumption of most Gaussian models. The use of the coefficient of determination (R^2) GOF for Gaussian models is likely very inappropriate for assessment or comparison of Poisson models. Appropriate GOF for Poisson models should be used. In Samet et al. (1997), appropriate GOF indices include the Poisson residual deviance, the overdispersion index (residual deviance/mean), and the Akaike Information Criterion (AIC). Gamble and Lewis have selected an inappropriate GOF criterion.

25. G&L P 843, C 1, L 36-46: "These limited data indicate PM is not of practical significance in explaining variability of mortality or morbidity. ... If one cannot measure the effect of a suspected risk factor, it is not logical to assert a cause-effect relationship."

RESPONSE: People get sick and die from many natural causes, and most of the variability in mortality or morbidity counts is due to the intrinsic variability of these natural processes. If the mortality or morbidity counts have a Poisson distribution, then the standard deviation of the number of events, as an index of variability, is equal to the square root of the mean or expected number. For London, with about 300 winter deaths per day, the standard deviation is about 17, or 5.8% of the mean; for Steubenville or Utah Valley, with about 3 deaths per day, the standard deviation is 1.73 or about 58% of the mean. Variability is even larger with hyper-Poisson data, as in many of the Philadelphia TSP studies. In most of the recent studies, most of the variability in health effects cannot be attributed to environmental factors such as weather or air pollution. There are some exceptions, such as the large increase in mortality and hospital admissions for heat-related symptoms in the July, 1995 Chicago heat wave, and the large increase in mortality and hospital admissions for cardiovascular and respiratory-related symptoms in the 1952 London air pollution episode. Rare high-level events provide evidence of causality. While air pollution events of such great magnitude no longer occur in the U.S. or in many other developed countries, there is reason to believe that adverse health effects still occur in susceptible American and European sub-populations at PM levels below the NAAQS.

26. G&L P 843, C 1, L 47-53: "The very low predictive power of PM ... increases the possibility that incomplete adjustment for confounding variables ... could result in a consistently [*sic*] small but spurious risk ratio."

RESPONSE: The issue of confounding is discussed at length elsewhere in these responses, and in much greater length in the PM CD. Studies whose results EPA finds credible have been analyzed by a number of different models specifically designed to evaluate potential confounders, and often by different investigators. After adjustment for seasonal and long-term time trends, weather effects can be adjusted by any of several strategies. Good-fitting PM moving averages almost always include concentrations for 1 and 2 days prior to the response day.

Copollutants can be confounders, but even in Philadelphia TSP studies, the TSP effect can be separated from other pollutants, except possibly SO_2 (Samet et al., 1996a,b; EPA, 1996). In other studies, such as the study in Santiago, Chile by Ostro et al. (not cited by Gamble and Lewis, but discussed and evaluated in detail in the PM CD), estimated PM_{10} effects can actually increase when the model includes copollutants.

The most convincing evidence is that different studies with different model-building strategies have found small but quantitatively similar PM effects estimates in cities in which many of the potential confounding factors are nearly absent: cities with hot climates and cities with cold climates, cities with dry climates and cities with wet climates, cities in which SO_2 levels are high and cities in which SO_2 levels are low. Gamble and Lewis find a causal hypothesis implausible for similar findings among many of the studies EPA cited. EPA finds much less plausibility in the argument advanced by some commenters that hypothetical, but as of yet unidentified and unsubstantiated, confounders account for PM associations in each and every study: if SO_2 is absent, then attribute the health effects to O_3; if SO_2 is absent and O_3 levels are low, then attribute the effect to CO or NO_2; if the PM effect is significant even when high levels of SO_2 and O_3 occur, then attribute the effect to humidity; and so on.

Small RR does not imply a negligible health effect. For example, RR = 1.05 per 50 $\mu g/m^3$ PM_{10} found in some studies is equivalent to an increase of 10% in mortality rate between days with 50 $\mu g/m^3$ and the same days with 150 $\mu g/m^3$ PM_{10}. Bradford Hill (1) notes that "we must not be too ready to dismiss a cause and effect hypothesis merely on the grounds that the observed association appears to be slight. There are many occasions in medicine when this is in truth so." Indeed, it is fortunate that the magnitude of the association is not "large"; RR above 1.5 - 2.0 for these studies noted as being needed to be "strong" enough to be considered strong evidence for causality would imply massive numbers of deaths in urban populations due to current ambient PM exposures. Relative risks of the size (RR > 1.5 or 2.0) argued as needed to support a finding of causality of PM effects were in fact, observed during earlier severe air pollution episodes, as noted earlier (e.g., for 2 weeks following famous 1952 London fog, RR = 2.6 for all cause mortality, RR = 9.3 for bronchitis, and RR = 2.0 for hospital admissions).

Gamble and Lewis also appear to ignore several epidemiological studies conducted at low PM concentrations in U.S. and European cities, including both short- and long-term exposures to PM air pollution, that find statistically significant relative risks of respiratory symptom categories in children in the range of 1.5 to 5 (Schwartz et al., 1994; Pope and Dockery, 1992; Braun-Fahrlander et al., 1992; Dockery et al., 1989; Dockery et al., 1996). Concentrations in these studies extend from moderately above to well below those permitted by the current PM_{10} standards. While, as noted by Gamble and Lewis, most of the recent epidemiological studies of mortality and hospital admissions report comparatively small relative risks, the findings of relative risks well in excess of 2 for earlier studies of high PM episodes, as well as the relative risks of 1.5 to 5 reported in more recent studies of less serious, but still important effects categories, lend credibility to EPA's interpretation of the results with respect to causality.

27. G&L P 843, C 1, L 54 to P 844, C 1, L 17: "**Specificity** ... **Against**. Reanalysis of data from Philadelphia by Li and Roth (10) do not show ... Styer et al. (3) did not find ... Lyon et al. (4) found that in Utah County the association was strongest for CVD (RR = 1.13) ... Inconsistent with Pope et al. (5) in Utah County, where the associations with both respiratory disease and CVD were stronger than for total mortality (RR = 1.20 and 1.09 vs. 1.08). ..."

RESPONSE: Gamble and Lewis do not offer much discussion on specificity - they note a few purported inconsistencies in the data sets. Hill (1) might actually find the higher relative risk for mortality for respiratory and CVD as specific for PM exposure. Medical evidence during mid-century high particle pollution episodes focus almost exclusively on these causes. Even if there were some misdiagnosed cases, no other plausible diagnoses were evident.

The Philadelphia TSP studies were also evaluated by other investigators. The HEI cause-specific Poisson regression analyses identify strong associations of CVD mortality with both TSP and SO_2; they find a significant relationship of pneumonia (i.e. respiratory) mortality to TSP alone, but only marginally significant ($t = 1.8$) when SO_2 is included; and virtually no relationship of TSP to cancer mortality (Samet et al., 1995). In Sec. 12.6 of the PM CD some methodological issues in (10) are briefly discussed. EPA believes that results from the methodologically appropriate investigations by HEI are more credible.

Cause-specific mortality in Utah Valley was subsequently reanalyzed by Pope and Kalkstein (1996). A large number of alternative models were evaluated, with reasonably similar findings of strong statistically significant cardiovascular effects. The estimated pulmonary mortality effects were larger in magnitude, but in some models were not statistically significant. There was modest sensitivity of RR estimates to weather and time trend model specification, but no qualitatively different conclusions. The less robust statistical significance of the pulmonary mortality risks may reflect the fact the cardiovascular deaths are about half the total number, whereas the pulmonary deaths are a much smaller fraction, less than 10 percent of the total.

The Lyon et al. Utah Valley study (4) may have obtained different findings for respiratory deaths because of the lower power of their analyses to find these effects. Their estimate of cardiovascular RR cannot be easily compared with Pope et al. (5) or with Pope and Kalkstein 1996) because the PM index they used was a dichotomous variable.

The Salt Lake City and Chicago study by Styler et al. (3) was briefly discussed in the PM CD. The analytical approach used by these investigators may have greatly reduced the power of the study to detect mortality effects, and further subsetting by cause of mortality further reduced the power. However, even with such reduced statistical power, Styer et al. still found significant PM-mortality associations in Chicago, a much larger city, when analyzed across the full year.

Detailed examination of the studies suggests that respiratory diagnoses are the most specific to PM air pollution, but are harder to detect than CVD because they occur much less often. The smaller RR for CVD may be a larger public health problem, and easier to detect in epidemiology

studies, because mortality diagnoses for CVD occur much more often than for respiratory mortality.

28. G&L P 844, C 1, L 18 to P 844, C 1, L 31: "**Coherence** ... Do the data 'conflict with the generally known facts of the natural history and biology of the disease'? ...mortality ... associations should also be observed with morbidity endpoints such as increased health care visits for respiratory illnesses, exacerbations of asthma, increased respiratory symptoms, and declines in lung function (30). ... hospital admissions are summarized in Table 2. "

RESPONSE: Hill (1) tells us that coherence means "our data should not seriously conflict with the generally known fact of the natural history and biology of the disease." Results of morbidity studies may not always be strongly related to the actual natural history leading to mortality. For example, individuals who are admitted to hospital for respiratory or CVD symptoms may be less likely to die than those who are not treated soon enough or aggressively enough. Individuals who visit physicians, clinics, or HMO's more frequently may receive better preventative measures and early interventions that in fact reduce hospital admissions and mortality. Thus, specific case histories may not provide enough evidence to be considered as indicators of coherence between symptoms, hospital admissions and mortality.

A more difficult issue is that the high-pollution events identified respiratory illness and CVD as related to PM exposure, but did not sufficiently clarify the relationship of specific respiratory symptoms to hospital admissions or deaths. The role of PM in the etiology and exacerbation of asthma is not known. Consequently, Gamble and Lewis did not indicate what specific morbidity precursors need to be found coherent with which other consequences in order to establish proof of coherence. For example, it is not clear that studies in asthmatic subjects cited by Gamble and Lewis are relevant to judging coherence with effects among the most susceptible adult subpopulations (e.g., the elderly, those with CVD, etc.).

Gamble and Lewis ignore much of the large hospital admissions literature cited in the PM CD. They selected one paper published in 1981 (15) which is reported as only showing association when TSP levels are >150 $\mu g/m^3$ as evidence that coherence does not exist here. EPA's reading of this paper suggests that Gamble and Lewis may have misread the purpose of the study, and it is puzzling that Gamble and Lewis were either unaware of or chose to ignore numerous other studies cited in the EPA PM CD as showing significant PM-hospital admissions associations.

29. G&L P 844, C 1, L 33-59: "A second guideline for coherence is whether the time-series morbidity data conflict with known facts about asthmatic admissions to hospitals ... "

RESPONSE: It is interesting to note that G&L selectively chose to cite asthma hospital admission increases while ambient PM levels decreased as an example indicating lack of coherence of the PM epidemiology findings with other facts. They did not cite the fact that cardiovascular mortality rates, in contrast, decreased during the same time period that PM concentrations were decreasing in the United States — a more pertinent fact in view of far

stronger evidence for PM effects on cardiovascular morbidity/mortality than for PM-asthma impacts. The Agency finds neither of the above two examples to be particularly cogent arguments for or against the coherence of PM effects. Obviously, just as one would not want to attribute the declining cardiovascular morbidity/mortality to PM reductions (given advances in heart medications, emergency treatments for heart attacks, etc.) over the same period of time, nor would it be correct to argue for lack of coherence based on increasing asthma rates in the face of nationwide PM reductions when other countervailing factors appear to be causing a rise in asthma rates.

30. G&L P 844, C 2, L 8-18: "**Against.** The PM/morbidity studies are of the same design as PM/mortality studies and are, therefore, subject to the same biases ... "

RESPONSE: As discussed in the preceding responses to comments on the mortality studies, daily morbidity time-series studies are designed to detect changes in morbidity counts in relation to changes in PM, weather, and copollutants over short time intervals. Personal exposure measurements do not appear necessary to characterize changes in exposure from regional air pollutants with outdoor sources, including PM and sulfates. Most of the recent morbidity studies have been subjected to model development strategies and specification searches similar to those of the mortality studies, so that the results of the recent time series morbidity studies are also believed to be reasonably robust with respect to the same potential confounders.

31. G&L P 844, C 2, L 26-30: "... in Steubenville, Samet et al. (15) reported no consistent associations of emergency room visits with TSP except when TSP > 150 $\mu g/m^3$ and temperature was high."

RESPONSE: See response #28 above. Reanalyses of the data in this 1981 study using more powerful statistical methods similar to recent morbidity time series studies would be interesting. The approach in (15), splitting a continuous E-R relationship by arbitrary cutpoints, almost certainly reduced the power of the study to detect effects. Even so, their finding of a larger TSP morbidity effect on hot days is consistent with the findings in a Philadelphia TSP mortality study (36).

32. G&L P 844, C 2, L 49 to C 3, L 4: "The one time-series morbidity study with personal measures of PM exposure ... (57) ... is not coherent with time-series ecologic studies of hospital admissions."

RESPONSE: Gamble and Lewis never indicate superficially the relevance of the referenced (57) morbidity study of asthma admissions to judging coherence of PM studies showing effects mainly in elderly patients (aged 65+ yrs) with preexisting cardiovascular and/or respiratory diseases, e.g., COPD or pneumonia. Moreover, the relatively small number of subjects and study details limit the comparisons that can be made even with other studies of asthmatics or hospital admissions. More specifically, Gamble and Lewis fail to note the key conclusions stated by the authors of the study (57);

> "The difficulty in establishing an effect of particulate matter on asthmatics' lung function has been identified in this analysis, i.e., the confounding effect of drug use masked the negative effect of particulate matter exposure in asthmatics. However, the data suggest that drug intake by patients who participated in the study ameliorated the potential negative effects of particulate exposure (Silverman et al., 1992)."

The main thing shown by the study is that the asthma medications work as expected, likely obscuring the full magnitude of PM effects hinted at by analyses corrected for season and mediation use. Perhaps a more meaningful analysis would have been to evaluate whether asthma medication use increased as PM or other air pollution variables increased.

33. G&L P 844, C 3, L 5-17: "Chamber studies have used exposure mixtures of polluted air, acid aerosols, and environmental tobacco smoke (ETS), agents that are similar to portions of ambient PM."

RESPONSE: Gamble and Lewis fail to note a number of limitations inherent in chamber studies published at the time the PM CD was completed that call into question the kind of inferences Gamble and Lewis attempt to base on such studies with respect to coherence and biological plausibility. Such limited studies as existed at that time can neither confirm nor refute the available epidemiological data. Clearly, chamber studies of sensitive humans at concentrations presenting a significant risk of mortality cannot be contemplated, although such exposures do occur in the ambient environment. More specifically, the controlled human studies referred to by Gamble and Lewis involving asthmatics or otherwise healthy subjects typically have employed only relatively brief (< 3 hr) exposures of very limited numbers of subjects to various PM agents, whereas the recent PM time-series epidemiology studies generally find that effects occur after longer exposure periods and might require a cumulative exposure, with the largest effects following elevated PM exposures of 1 to 3 days duration, not just a few hours. The relative risk of effect in the hospital admission and mortality studies suggest that, even at relatively high concentrations in chambers, single exposures would require on the order of thousands or more subjects to detect significant effects. It is therefore important in the design of such studies to examine more sensitive intermediate endpoints that might permit detection with more feasible sample sizes. However, chamber studies have generally focused on measurements of lung functions and symptoms, which may not be related to hypothesized mechanisms involving inflammation, mediation through neural pathways, and effects on the cardiovascular system.

The available animal toxicological data are similarly limited. In order to provide further perspective on the difficulties in comparing the laboratory data in the PM CD with the epidemiological studies, it is appropriate to draw excerpts from a summary of these issues presented in Appendix D of the staff paper:

EXTRAPOLATION OF RESULTS FROM LABORATORY STUDIES TO THOSE OF EPIDEMIOLOGIC STUDIES: STRENGTH AND LIMITATIONS OF CONTROLLED HUMAN AND ANIMAL STUDIES

The adverse effects of particulate matter exposure have been shown to be consistent between historical and more recent studies. The effects can be severe and tend to be concentrated in sensitive sub-populations who have pre-existing conditions or characteristics that tend to make them vulnerable to respiratory insult (the very young and old, asthmatics, COPD patients, patients with pneumonia etc). The additional risk of reported mortality and morbidity from particulate matter exposure is relatively small in terms of the whole population. Therefore, large numbers of people must be exposed before effects can be discerned in studies. The question arises as to how to elucidate the mechanism of action of particulate matter in humans. What are the considerations that must be taken into account when an analysis of the body of human clinical data and experimental animal work is done in order to infer a plausible mechanism for particulate matter effects?

a. Numbers of Individuals Affected

An issue of primary concern is that of statistical power. The nature of the effect described in epidemiological work is consistent, and serious, but occurring in a relatively small fraction of the total population (e.g. 1 in a million increased risk for daily mortality). Therefore, theoretically a relatively large number of animals would be needed to mimic the frequency of response at similar doses. The use of a similar number of animals to mimic the frequency of response to ambient air concentrations of particles which have been associated with effect in humans is impractical. Therefore, in many experimental paradigms, relatively large concentrations are often given investigate the response from a limited number of animals. However, the questionable relevancy and sensitivity of such paradigms limits their use in the determination of the mechanism of action of relatively low changes in concentrations of inhaled particulate matter.

b. Heterogeneity of Human Population

The human population for which the effects are most demonstrable are a sub-population from a genetically heterogenous group. Furthermore, consistency of response is highly variable among the population at risk (e.g., a relatively small group of asthmatics have aggravation of symptoms and not all patients with pneumonia or COPD die as a result of an increase in inhaled particle concentration). The CD suggests that for clinical studies involving asthmatics, differences among subjects may explain in part the differing results between laboratories who study effects of acid aerosols. As an example of differential susceptibility to a respiratory insult, a minority of individuals (3-5%) who are exposed to etiologic agents responsible for hypersensitivity pneumonitis (allergic alveolitis) will develop disease. Determinants of susceptibility for that disease have been

described as both the genetic constitution of the individual and the presence of preexisting lung disease. Similar factors probably play a role in susceptibility to inhaled particulate matter effects.

By contrast experimental animals are bred as much as possible to be homogenous genetically so as to give great consistency in response. They are also usually studied in their prime in regard to age and general health. Presence of disease is generally considered to be a confounding factor to be stringently controlled in most animal paradigms. As stated above, those segments of the general population most affected from PM_{10} exposure are the sick, the very young, and the old. Therefore the sensitivity of studies using relatively small numbers of healthy, genetically homogenous, laboratory animals who are in their prime is diminished in exploring mechanism of particulate matter effects.

c. Heterogeneity of PM_{10} Composition

Another key element helps to frame the discussion of the relevance of human clinical studies and experimental animal work to establish a mechanism of action of particulate matter in humans. That is the issue of heterogeneity of both the composition of and exposure to particulate matter. Particulate matter is a broad class of physically and chemically diverse substances (as described in Chapter IV). The PM_{10} fraction is composed of two distinct sub-fraction of particle: fine and coarse particles. Additionally, the properties of PM_{10} vary greatly from place to place because of differences in source mixes and atmospheric conditions.

Thus unlike a typical experimental paradigm, where the agent to be studied is isolated and the effects of exposure described in well controlled studies, the heterogeneity of the PM_{10} entity forces a different experimental approach. Typically constituents of the fraction are tested individually to see if effects similar to those observed in humans are reproduced. Consequently, animal studies are further weakened in regard to ability to establish a mechanism of action of particulate matter and to either refute or validate epidemiological observation of effect in humans.

d. Dosimetric Heterogeneity

Finally, dosimetric comparisons between laboratory animals and humans, show that there are significant differences in the respiratory architecture and ventilation of the two which adds additional complication to comparisons of experimental and observed data. Ventilation differences coupled with differences in upper airway respiratory tract structure and size, branching pattern, and structure of the lower respiratory tract occur between species as well as between healthy versus diseased states. These differences may result in significantly different patterns of airflow affecting particle deposition patterns in the respiratory tract (CD, Chapter 13). Additionally, inter-species variability

in regard to cell morphology, numbers, types, distribution, and functional capabilities between animal and human respiratory tracks, leads to differences in clearance of deposited particles which may in turn affect the potential for toxicity (CD, Chapter 13). Consequently the difficulty of using experimental animal data to investigate particulate matter effects is further defined.

e. Lack of Distinct Disease Pathology

The background levels of cardiopulmonary disease as the cause of death for the general population is very high. Given that COPD and heart diseases are frequent causes of death, it is difficult to discern those who die from the additional effects of particulate matter from those already dying from such diseases and to do autopsy to identify a specific pathology associated with particulate matter caused mortality. Even in historical studies involving higher levels resulting in more pronounced effect it is hard to get an adequate characterization of pathology related to particulate matter effects. Thus without such a characterization of the pathology of particulate matter induced mortality, development and validation of appropriate models to study such effects are more difficult.

f. Lack of Appropriate Equivalents to Epidemiological Endpoints

Animal toxicological equivalents of such epidemiological endpoints as hospital admissions and emergency room visits as an indication of morbidity cannot be obtained. Although mortality can be recreated in a laboratory setting, the relevance of mechanism is currently an issue. In addition, there is question as to what the most appropriate measure of particulate matter is in regard to its toxicity.

In addition to considerations of dose (inhalability and appropriate metric), the nature of the response to particles and correlations of the appropriate response to susceptible population are yet to be resolved. Thus, identification of the dosimeter which induces mortality and morbidity has not been elucidated with consequent difficulty interpretation and design of controlled animal and human studies. [U.S. EPA, 1996b; Appendix D]

34. G&L P 844, C 3, L 28-38: "Bauer et al. (61) exposed 11 elderly patients to COPD to 75 $\mu g/m^3$ H_2SO_4 for 2 hr ... The lack of airway obstruction and shortness of breath on exercise ..."

RESPONSE: This brief exposure to a small number of subjects suffers from most of the difficulties noted above. Despite using representative sensitive subjects, the exposure was unlikely to have adequately mimicked "real world" exposures. Cumulative exposures for at least 24 hours may be necessary. Moreover, based on the epidemiological results, no significant serious effects would be expected in such an extremely small number of individuals, whose susceptibility in terms of key responses was undefined. Also, the respiratory mechanical

endpoints tested may not have been the most relevant. Bauer et al. did not include effects on clearance, or indices of inflammation, or cardiovascular responses that appear to be germane to current hypothesis regarding potential mechanisms of action. For these reasons, these results are inconclusive and cannot be used, as Gamble and Lewis suggest, to suggest a lack of coherence with the epidemiological results.

35. G&L P 844, C 3, L 52 to P 845, C 1, L 18: "There are several chamber studies of healthy and asthmatic volunteers exposed to ETS containing PM ... for 2 hr ... for 4 hr. The reductions in FEV_1 were completely reversible and were not severe enough to result in hospital admission."

RESPONSE*:* Again Gamble and Lewis do not explain how the ETS exposure studies, per se, are relevant for judging the coherence of the findings from the ambient PM epidemiology studies with respect to mortality in sensitive subgroups. The atopic asthmatic subjects in the studies cited may, however, be relevant to the less prominent issue in the PM literature of aggravation of asthma. Given that Gamble and Lewis believe ETS has relevance to outdoor combustion particles, it is surprising that they do not note the apparent coherence of these findings with the epidemiological results suggesting ambient PM contributes to aggravation of asthma.

By selection criteria, the atopic asthmatics in the studies (65,66) were sensitive to the effects of smoke. Brief (2-4 hour) exposures to levels of 850 $\mu g/m^3$ resulted in 10% of those exposed responded with FEV_1 changes of greater than or equal to 20%, a benchmark that has been used to define an adverse health effect in the case of ozone. In the face of these results, Gamble and Lewis point out that no one was hospitalized and the effects were reversible. The former outcome was, of course, required by medical ethics to be ensured by the study design. Only the least sensitive of the group were exposed to the still higher levels Gamble and Lewis also cite. Establishing the ETS tolerance of the least sensitive asthmatics does nothing to address the question of whether a greater response might occur in the most sensitive asthmatics that might be found in a distribution of the much larger population that occurs in cities. In drawing negative conclusions from these studies, Gamble and Lewis do not suggest how they were able to compare the relationship between a 10% incidence of adverse lung function changes from 2- 4 hour exposures of about 20 to 30 resting asthmatic subjects in these controlled studies to the results of 24 hour or longer exposures to a complex of PM pollutants for the entire asthmatic population of a large city at varying activities, but at lower concentrations (i.e. Thurston et al., 1994). Clearly, it is inappropriate to conclude that the chamber studies have established a no-adverse effects level for ETS, much less ambient fine particles. Given the variation in activity levels, PM composition, peak to mean ratios for outdoor pollution, and sensitivity, these controlled studies may well support the plausibility of observed associations between PM and aggravation of asthma. Contrary to Gamble and Lewis assertions, consideration of these studies more readily supports, not contradicts, coherence of PM studies, at least with respect to effects on asthmatics.

36. G&L P 845, C 1, L 19 to C 2: "**Biological Plausibility.**"

RESPONSE: With respect to biological plausibility, Hill noted that "this is a feature I am convinced we cannot demand. What is biologically plausible depends upon the biological knowledge of the day" (Hill, 1965). This statement is clearly pertinent to the toxicological and mechanistic understanding of the effects of PM and associated air pollutants, especially at lower concentrations. The mechanistic evidence published as of the time the Criteria Document closed does not provide quantitative support for the epidemiological results, but neither can such limited evidence refute these findings. It is also important to stress that our understanding of biological mechanisms for PM pollution effects is not sufficient to explain the effects observed at much higher concentrations in historical air pollution episodes, for which causality is generally accepted. Moreover, the toxicological literature has only recently begun to examine animal models (or controlled human studies) that might reflect the sensitive populations in question (the elderly, individuals with chronic respiratory and cardiovascular disease) or that adequately reproduce all of the physico-chemical properties of particles in the ambient atmosphere. In short, the absence of evidence of a particular mechanism is hardly proof that there are no mechanisms that could explain the effects observed so consistently in the epidemiological studies.

A supplemental observation not part of the above response is worth noting. Several recent papers, which appeared too late to be examined by Gamble and Lewis or to be considered in EPA's review of the criteria and standards, have reported findings that add support for biological plausibility of the PM epidemiology findings and/or advance testable hypotheses of plausible mechanisms by which low level PM effects (mortality, morbidity) could be induced. Although not considered in the final promulgation decision, some of these are provisionally summarized in Appendix C.

37. G&L P 845, C2 to P 846, C1, L 17: "**Bias**. ... **Ecological Study Design and the Ecologic Fallacy Bias**. ... **Against**. Epidemiology text books and articles ... note that ecological studies are limited ... based on group data (42, 44-46)."

RESPONSE: Gamble and Lewis have ignored the fact that the primary feature of daily time-series studies -- the ability to do longitudinal comparisons in a virtually unchanging population -- avoids some of the major perceived difficulties in true "ecological" analyses. Daily time series studies of mortality, hospital admissions, and other health effects involve observations of effects in populations rather than in individuals. The predictors of effects include observations of air pollution and weather at one or a few locations in the SMSA, rather than for each individual. In this sense -- and only in this sense -- these studies may be described as "ecologic". The specific character of these time series analyses allow for much greater control of potential confounding factors than is commonly seen in other population-based study designs, e.g. in cross-sectional ecological studies across different SMSA's, such as described in Section 12.4.1.2 of the 1996 PM CD (see also preceding EPA responses made above to G&L comments).

The main advantage of daily time series studies relative to other so-called ecologic study designs is that it is possible to examine *differences of health effect in response to differences in air*

pollution exposure in essentially the same population over short periods of time. Air pollution concentrations typically change over periods of a few days, with high concentrations of regional air pollutants such as PM typically associated with the temporary formation of low-altitude atmospheric inversion layers, that is, with air pollution episodes. During this same period of time and for some period of time before and after the high-concentration episodes, there are likely to be similar weather conditions with respect to temperature, humidity and other meteorological variables, but with much lower air pollution concentrations. During a period of a few weeks or possibly even a few months, the temporal changes in air pollution can be correlated with temporal changes in health effects in nearly the same population (or population sub-group) and over a generally similar range of weather conditions. These effects may then be statistically aggregated over many years. Daily time series study designs therefore allow for a much higher degree of adjustment for potential confounders than do other population-based study designs.

In summary, time series studies have distinct advantages in study design over some other epidemiology study designs that Gamble and Lewis class as "ecologic studies". While the commenter has assembled some textbook-type general comments about ecologic study designs, EPA does not find these comments to be particularly relevant to the specific time series studies cited in the PM CD.

38. G&L P 846, C 1, L 20-40: "The logical fallacy in the time-series studies is that the concentration of PM collected from a sampler in a metropolitan area is not a reasonable proxy for personal exposure to PM. The limited data available suggest that correlations of ambient concentrations with indoor and personal exposures are generally close to zero. ... Spengler et al. (77) conclude that ambient measurements were poor predictors of personal exposure and that ETS is the dominant source of indoor air pollution."

RESPONSE: Gamble and Lewis miss the mark on the key issue involving indoor/outdoor and personal exposure issues. The key question is not whether the central monitoring site measurements contain a signal reflecting actual exposure to total PM from all sources at the individual level. Rather, the ultimate question is whether the central monitoring site measurements contain a signal reflecting average population exposure to ambient PM, including both exposure to ambient PM while outdoors and to PM of ambient origin that has infiltrated indoors. The proposed PM standard is intended to protect the public from exposure to ambient PM. It is not intended to protect them from PM generated in their homes or by their personal activities. There is ample evidence, as discussed in Chapter 7 of the PM CD that personal exposure to PM of ambient origin, while outdoors and while in indoor micro-environments, does correlate with concentrations measured at central monitoring sites.

Central site ambient PM concentrations are not correlated with exposure to PM generated indoors, or by personal activities while outdoors such as smoking. Therefore, such latter types of exposures would not be a confounder in epidemiologic studies of ambient PM effects. They could, however, be independent risk factors.

EPA therefore concludes that indeed there may not always be a significant correlation between ambient concentrations and personal exposure to total PM from all sources. However, this information is irrelevant to EPA's regulation of ambient PM. The significant factors are: (1) there is a relationship between health outcomes and ambient PM concentrations; and, (2) there is a relationship between ambient PM concentrations and personal exposure to PM of ambient origin.

The issue of relationships between ambient PM and personal exposures, including indoor, to PM of ambient origin is discussed in more detail elsewhere (See Appendix D).

EPA also concludes that it is reasonable to presume that reductions in ambient PM concentrations will reduce personal exposures to PM of ambient origin and will protect the public from adverse health outcomes associated with personal exposure to ambient PM.

39. G&L P 846, C 1, L 54 to C 2, L 7: "The range of opinion regarding ecologic studies ..."

RESPONSE: These opinions are irrelevant. See Response #37 above (for G&L. P 845, C2 to P 846, C1, L 17.

40. G&L P 846, C 2, L 8-16: "The differences between actual and measured exposure are rarely determined. Therefore, measurement error for all the independent variables in time-series ... results are complex and not always predictable."

RESPONSE: If the purpose of the time-series analyses is to evaluate underlying relationships and determine the "true" relationship, then it may or may not be useful to evaluate measurement error models, depending on the magnitude of biases. If the purpose of the analyses is to develop predictive models, then this comment is irrelevant. For purposes of prediction (e.g., as in risk assessment and regulatory impact analyses), when the input for the model predictions has the same measurement error distribution as the data that were used to estimate the model parameters, then the parameters of the fitted model may be used for prediction without further adjustment.

Also see response to issue 22 above for further elaboration on the effect of measurement error in PM studies.

41. G&L P 846, C 2, L 17-40: "Lipfert and Wyzga (82) performed data simulations and numerical experiments using mortality and pollutant data from Philadelphia ... In a standard multiple regression procedure, a pollutant having a lower measurement error yields an inflated coefficient or risk estimate, while the coefficient of the pollutant with the higher measurement error is reduced. ... may provide a partial explanation of why SO_2 but not PM effects are often attenuated in regression analyses."

RESPONSE: This phenomenon may occur mathematically, but there is as yet little empirical evidence to evaluate the magnitude of the biases, which may be negligible. The basis for the

findings in (82) can be understood without extensive mathematical analyses or computer simulations. Suppose that ambient air pollutants called A and B have highly correlated true concentrations, but A is a good measure of actual ambient exposure, while B is measured with a large amount of error in ambient exposure characterization or with large analytical error. Because they are strongly correlated, either A or B can serve as a partial surrogate for the other pollutant. However, since A is more accurately measured than B, the observed value of A with small error will better predict the response than will the poorly measured surrogate B. To the extent that one variable can be used to replace the other, one will obtain a much better estimate of the response by attributing B's effects to A than by attributing A's effect to B. Since A and B are assumed to be correlated, any statistical procedure based on optimizing goodness of fit of the model to the data, including multiple linear regression and Poisson regression models, will select the better measured variable (i.e. will inflate the coefficient of A in the model) because A is a better predictor of B's effects than is the poorly measured value of B itself.

EPA examined the statistical arguments advanced in the study by Lipfert and Wyzga (82), and found a limited set of statistical conditions under which measurement errors can actually result in inflated effects estimates and relative risks. This analysis, described in Appendix D, finds that in order for measurement errors in one pollutant variable to significantly bias the estimated effect of another pollutant, three conditions are necessary: (1) the measurement error in the poorly measured pollutant must be very large, roughly at least the same size as the population variability in that pollutant; (2) the poorly measured pollutant must be highly correlated with the other pollutant, either positively or negatively; and,(3) the measurement errors for the two pollutants must be highly *negatively* correlated. This important factor was not considered in Lipfert and Wyzga (1995) or by Gamble and Lewis. These authors have provided no evidence suggesting that such conditions are likely to occur with respect to the measurement of ambient PM in relation to those for gaseous co-pollutants commonly used in epidemiological studies.

As discussed in the response to issue 22 above and in Appendix D, it therefore appears unlikely that measurement and exposure errors for PM and other pollutants have inflated the estimated effects of PM, even in multivariate analyses. More importantly, the available evidence on the consistency of the PM-effects relationships in multiple urban locations, with widely varying indoor/outdoor conditions and a variety of monitoring approaches, makes it less likely that the observed associations of PM with serious health effects at levels allowed under the current NAAQS are an artifact of errors in measurement of pollution or of exposure (U.S. EPA 1996b, pp. V-39 to -43).

EPA believes it is more likely that the SO_2 coefficient is not differentially attenuated because of measurement error, but is in reality low because SO_2 alone is less likely to reach the most important respiratory targets. As noted above, the Criteria Document assessment of the Philadelphia studies finds that PM can reasonably be distinguished from potential effects of all pollutants except SO_2. The Staff Paper builds on this analysis through an integrated assessment that draws on information from atmospheric chemistry, human exposure studies, and respiratory tract penetration results to provide insight as to which of these two pollutants is more likely to be

responsible for mortality in the elderly and individuals with cardiopulmonary disease (U.S. EPA 1996b; pp. V-46 to -50). That assessment notes that the inhalable (PM_{10}), including the fine ($PM_{2.5}$), components of TSP are more likely than SO_2 to penetrate and remain indoors where the sensitive population resides most of the time. In addition, these PM components, especially $PM_{2.5}$, penetrate far more effectively to the airways and gas exchange regions of the lung than does SO_2. Furthermore, in Philadelphia, it is possible that SO_2 is a surrogate for fine particulate acid sulfates. For these reasons, even though statistical analyses of the Philadelphia data set cannot fully distinguish between these two highly correlated pollutants, EPA believes that the weight of the available evidence from an integrated assessment more strongly supports the notion that PM is playing an important direct role in the observed mortality effects associations in Philadelphia. Moreover, as noted above, in some other locations with significant PM-mortality associations, ambient SO_2 levels are too low to confound PM.

42. G&L P 846, C 2, L 55 to C 3, L 21: "**Measurement validity bias.** ... Thus, the pollutant with the lowest measurement error will have the spuriously highest regression coefficient. ..."

RESPONSE: See Response to #22 and #41 above.

43. G&L P 846, C 3, L 22-49: "**Averaging time or lag bias.** ... the appropriate lag period must be used for the independent variables (82). The appropriate lag period is different for each pollutant in the model [and in different studies] ... If SO_2 and O_3 effects are related to peak exposures ... rather than 24-hr or longer means ... The variable number of lag times used in various studies also raises concern about the lack of rationale for a 24-hr averaging time for PM_{10}."

RESPONSE: There may be real biologically based differences in appropriate averaging times, depending on the circumstances. While respiratory and cardiovascular health effects have been identified as responsive to PM exposure over the preceding few days, these are broad categories of health effect. Some health effects may be more responsive to exposures on the preceding day or even the same day, while other effects may be responsive to exposures 2 or 3 or more days earlier. Since the baseline incidence of these effects depends on population susceptibility and on weather, which will surely differ from one region to another, it is hardly surprising that there is some variability among the *relative* risk findings of studies in different places or covering different time intervals, and differences in lag structure associated with these findings. The same could be said of other pollutants, whose effects may also involve different lag structures in different locations.

Some studies, however, have found significant PM effects associated with exposure during a single 24-hour period. Standards for PM concentrations during a 24-hr hour period will tend to ensure control to the same PM level during any daily moving average period, or any other 24-hr lag exposure period. Moreover, as discussed in the Staff Paper, annual average standards can reduce exposures for single or multiple days. This is a more robust way of setting national standards when dealing with the possibility that there are real differences in time lag PM effects

in different places. Such potential differences are, nevertheless, of interest and represent an excellent topic for future research.

44. G&L P 846, C 3, L 50-61: "The lags for weather ... in the summer is less than or equal to 1 day and is much longer in the winter. Three-day lag times were common in a number of U.S. cities ..."

RESPONSE: Appropriate seasonal adjustments for weather lag times is clearly a good idea. However, in studies where weather effects are adjusted over intervals of exposure by use of synoptic climatological categories, or by nonparametric smoothing of weather variables, this is very likely to be adequately adjusted. Lag structures for PM or other pollutants are probably correlated with lag structures for weather to some extent, since concentrations are correlated with weather variables. EPA regards methodologies used in recent PM studies as appropriate to deal with these issues.

45. G&L P 846, C 3, L 62 to P 847, C 1, L 15: "**Response function bias.** ... appropriate form of the exposure-response function ... Most models use a linear form even if more than one pollutant is in the model ..."

RESPONSE: Chapter 12 of the PM CD contained substantial discussions on the shape of the exposure-response relationship and the appropriate functional specifications for copollutants and for weather variables. There was a substantial discussion of the HEI results (Samet et al., 1995) on the joint response of Philadelphia mortality to TSP and SO_2 by season, with three-dimensional plots and contour surfaces. There was some indication of weak nonlinearity in this surface. Other univariate E-R relationships were plotted in Sec. 12.3 and 12.6. There was some indication of possible threshold effects, which was explored in detail. However, in a number of studies, formal significance tests of nonparametric smoothing models for the PM relationship vs. a linear model were carried out, and the linearity hypothesis was not rejected. EPA believes that these model specification tests are an extremely desirable component of model evaluation, and that visual as well as quantitative diagnostics for model fit should be presented whenever possible. However, in the absence of such diagnostics, linear models for PM and copollutant health effects (generalized linear model Poisson regression, log link) are reasonable and appropriate assumptions.

Weather models are clearly nonlinear and should be carefully tested for goodness of fit before PM and copollutant effects are evaluated, since weather is usually an important confounder of air pollution. Some obviously misspecified weather models have been proposed, such as a V-shaped E-R function for temperature (Li and Roth), when nonparametric smoothers suggest U-shaped functions. Many recent studies avoid model misspecification for confounders, such as weather and time trends, by using nonparametric smoothing functions.

46. G&L P 847, C 1, L 16-26: "**Horse blinder or tunnel vision bias**. Many of the studies have focused on PM only as the pollutant of choice, and have ignored other pollutants ..."

RESPONSE: First of all, model specification searches that include PM are based on the clear hypothesis that PM is likely to cause adverse health effects at PM levels below the current NAAQS. This hypothesis is well-justified by definitive results from the London mortality studies described in the 1986 PM/SO_x CD. Many of the more recent studies evaluated in Ch. 12 of the PM CD have, in fact, included additional pollutants in their analyses or reanalyses. EPA also assessed and compared study results with and without adjustment for other pollutants.

For example, EPA reanalyses of the Samet et al. (1996a,b) report on multiple pollutants in Philadelphia suggested that there was a general factor (principal component) involving TSP, SO_2, NO_2, and CO. There was also considerable separation of O_3 from this component except in summer, and another distinct component composed primarily of TSP and SO_2. The TSP coefficient remained statistically significant, even with the inclusion in the model of one or more other pollutants found in the same component**.**

The assessment of potential confounding of a possible PM effect by co-pollutants has typically been based on the sensitivity of the estimated PM regression coefficient to inclusion of other pollutants. Not surprisingly, most studies found substantial sensitivity, typically a reduction of the PM coefficient although this has not always been the case. It is also not surprising, in view of the air pollutant correlations, that adding one or more correlated pollutants to a regression model can greatly increase the estimated uncertainty of the PM coefficient and thereby decrease its statistical significance *even when the estimated size of the PM effect is virtually unchanged*. This is not necessarily the best way to separate the effects of multiple correlated pollutants, but published studies do not give EPA many alternatives except comparison of PM regression coefficients and RR when different pollutants are included in addition to PM. All available information on confounding by other pollutants was evaluated by EPA.

47. G&L P 847, C 1, L 27-36: "**One-sided reference bias.** ... Consideration of arguments and data contrary to a cause-effect relationship must be addressed ..."

RESPONSE: EPA evaluated all then available published data, and reported in detail findings of "negative" effects, or non-significant positive effects, within each study as well as across studies. Studies with more serious methodological flaws or with inadequacies in the data received lower weight in EPA assessments than studies without such flaws. EPA did not uncritically accept any study, but evaluated each study for strengths, weaknesses, and divergent results from comparable studies. EPA's assessments were reviewed by external peer reviewers, and then by CASAC. Public presentations and comments were received, and all of these comments and reviews were evaluated for inclusion in subsequent revisions of the PM CD and Staff Paper. All of the studies EPA received were evaluated; none was omitted. Some of the studies reviewed by EPA received little comment because, in the opinion of EPA staff, CASAC, and many external reviewers, these studies added little to the assessment of the available information. No comparably detailed review of the individual studies for internal validity is evident in the paper by Gamble and Lewis, and their examination of the range of studies is far less inclusive than that in the PM CD.

48. G&L P 847, C 1, L 37 to C 2, L 17: "**Confounding from Weather and Other Pollutants. ...Against.** There are several reasons confounding from pollutants and weather may in part explain the observed associations."

RESPONSE: This issue was discussed at length in the EPA PM CD and Staff Paper, and in the preceding responses to issues raised by Gamble and Lewis.

49. G&L P 847, C 2, L 18-37: "The high R^2 values for weather in climate time-series studies and the low R^2 values in air pollution studies indicate that there has not been adequate control for weather. Weather (and not necessarily mean temperature or relative humidity as commonly used in air pollution) show relatively high R^2 values (see Table 6 ...). Kalkstein and Davis (85) ..."

RESPONSE: First of all, R^2 is not an appropriate index for comparing mortality count regression models, as discussed above (Response #25) . Since daily mortality counts should have a Poisson or hyper-Poisson distribution for all practical purposes, goodness-of-fit indices appropriate to the Poisson distribution (such as the overdispersion index) should have been used, not R^2.

Secondly, these data refer to a much earlier period of time. The data were scattered over 11 years: 1964-66, 1972-78, 1980. The period of time December 25-January 1 was excluded from each data set, so these data do not constitute contiguous time series, as do most of the air pollution time series. It is likely that "exposure" to "ambient weather" may have been greater during the earlier part of the period, especially during summers in the earlier part of the study period when air conditioning was less widely used in residences and in transportation than during the years of the more recent air pollution studies.

Thirdly, and most importantly, many of the variables that Gamble and Lewis have characterized as "weather" variables may in fact be surrogates for air pollution. In particular, for the eight cities listed in Table 6 (Birmingham, AL; Chicago, IL; Cincinnati, OH; Detroit, MI; Los Angeles, CA; New York, NY; Philadelphia, PA; St. Louis, MO), airport visibility at 3 A.M. or at 3 P.M. is listed as a statistically significant predictor of total summer mortality in 4/8 of these cities, as a statistically significant predictor of age 65+ summer mortality in 3/8 of the cities, and as a statistically significant predictor of total winter mortality in 3/7 of the cities. In the larger set of Tables in (85), airport visibility at 3 A.M. or at 3 P.M. is listed as a statistically significant predictor of total summer mortality in 12 of 25 cities (Table 5 in (85)), as a statistically significant predictor of age 65+ summer mortality in 12 of 25 cities (Table 6 in (85)), and as a statistically significant predictor of total winter mortality in 9 of 30 cities (Table 8 in (85)). Since airport visibility has been used as a surrogate for fine particle exposure in a number of studies, it may just as well be playing the role of a fine particle index in addition to (or instead of) a weather index in these studies.

Wind speed at 3 A.M. or 3 P.M. is listed as a statistically significant predictor of total summer mortality in 6/8 of the cities in Table 6, as a statistically significant predictor of age 65+ summer mortality in 4/8 of the cities, and as a statistically significant predictor of total winter mortality in

4/7 of the cities. Very low wind speeds often occur under inversion layers. Under these conditions, concentrations of PM as well as other air pollutants are likely to be elevated, so that wind speed may also be serving as a partial surrogate for air pollution. Most investigators have not used wind speed as a weather covariate in air pollution studies for this reason.

Table 6 does not prove the hypothesis put forward by Gamble and Lewis.

50. G&L P 847, C 3, L 4-11: "Results from St. Louis, Missouri ... suggested fluctuations in daily mortality were much more sensitive to stressful weather than high pollution levels. The most stressful weather synoptic category was associated with highest mortality ... and did not have high pollution concentration."

RESPONSE: This argument is consistent with both weather and pollution contributing to mortality. Since the synoptic category with the highest pollution concentrations was not the most stressful weather category, but wind speed was a significant predictor of total mortality in both summer and winter in St. Louis (Table 6), then one may conclude that the effects of wind speed acting as a surrogate for air pollution was at least partially correlated with excess mortality. It would be possible to interpret these results as an indication that the effects of weather and air pollution are separable. The recent HEI study (Samet et al., 1997) establishes the separability of weather effects and air pollution effects in Philadelphia much more directly.

51. G&L P 847, C 3, L 12-16: "The same model without PM or visibility explained 51% of the variability in elderly mortality. PM was not significant in these analyses. These results are contrary to the analysis of St. Louis by Dockery et al. (74)."

RESPONSE: The time periods of these studies did not overlap: 11 of the years from 1964 to 1980 were used in (89), and 1985 to 1986 were used in (74), so that the comparability of the results is doubtful. The study (89) used TSP as a PM index, and visibility (only a surrogate for FP), as well as wind speed and direction (significant confounding factors). The PM indicators in (89) are much less informative than the use of PM_{10} or $PM_{2.5}$ in (74). The analyses in (74) were verified by HEI (32). A recent detailed reanalysis of the St. Louis study by Schwartz, Dockery, and Neas (1996) using appropriate analytical methods and a much longer data base, 1979 to 1987, found even stronger statistically significant relationships of mortality with $PM_{2.5}$ or PM_{10} in St. Louis than did (32, 74). The more recent results with appropriate PM indicators are more credible.

52. G&L P 847, C 3, L 17-23: "It seems reasonable to infer from these data by Kalkstein (89) that weather is a stronger risk factor than PM; temperature and relative humidity do not adequately adjust for weather and are correlated with PM; and weather is confounding the PM/mortality association."

RESPONSE: G&L appear to have substantially misinterpreted these papers, and to have misused the results. The synoptic categories in (89) are substantially defined by differences in visibility,

wind speed, and wind direction, as well as by temperature and humidity. The highest TSP concentrations (mean values 101-103 $\mu g/m^3$) in St. Louis were found during weather Synoptic Categories (SC) 2, 3, and 6, with moderate SE or SSE winds. These SC's had low visibility (7-9 km in SC 2, 5-7 km in SC 3, and 9-11 km in SC 6). The lowest TSP levels (means of 70 to 76 $\mu g/m^3$) occurred during: SC 7, SC 8, SC 10; good visibility of 12-14, 9-11, and 12-14 km respectively; and moderate to strong W, WSW, and NW winds. These correlations of TSP, visibility, wind speed and direction must be considered as substantially confounded with SC. In some situations, such as the Utah Valley (Pope and Kalkstein, 1996), either the use of SC or nonparametric smoothers of temperature and humidity is adequate. Many of the recent studies used by EPA have used nonparametric smoothers of weather and time, which are believed to provide satisfactory adjustment for these factors without introducing unnecessary confounders such as visibility or wind speed.

53. G&L P 847, C 3, L 29 to P 848, C 1, L 7: "The lack of control is probably due to the use of improper metrics to measure weather and inappropriate lag times for temperature. ... The appropriate lag periods for hot temperature are 0-1 days; for cold temperatures, the lag periods are somewhat longer ... "

RESPONSE: The methods used in many air pollution studies were developed to provide a basis for attributing health effects to non-pollution factors, including weather, and should not necessarily be regarded as appropriate causal models for weather effects. These methods do appear to achieve their intended purposes: (1) empirical adjustment of air pollution health effects models for potential confounders; (2) reduction of unaccounted variability such as overdispersion in health effects models by accounting for other independent risk factors. The development of biologically-based models to adjust for weather effects is of scientific interest.

Appropriate lag structures for modeling the health effects of air pollutants have been assessed in most of the recent studies cited by EPA. Such evaluations are an indispensable part of evaluating models fitted to health effects data. There has been a range of findings on lags for PM and other air pollutants, and EPA cannot preclude the possibility that there may exist some real site-specific differences. Site-specific differences in moving averages or lag times may be related to differences in the chemical composition and size distribution of PM at different locations, to differences in other air pollutants in the atmosphere, to differences in climate or weather that may affect exposure, and to differences in baseline disease incidence in the population that may be affected by PM. The proposed 24-hr standards should allow an adequate margin of safety for such differences in the temporal pattern of response to PM exposure, however.

54. G&L P 848, C 1, L 27-35: "Copollutants ... have not been included in most air pollution studies ..."

RESPONSE: Studies on PM_{10} that EPA finds most credible are those studies in which copollutant effects have been assessed. In some of these studies, including copollutants may

attenuate the estimated PM effect to the point where it is no longer statistically significant. However, this has not happened in other studies. In general, the inclusion of copollutants produces a somewhat smaller PM effect size estimate, but usually remains positive and overall statistically significant in a meta-analytic framework. Because air pollution mixtures differ from one location to another, some variability in findings should be expected. The results suggest that analyses using only PM as a pollutant have some degree of robustness when other pollutants are also included, but that this will not be detected in all studies. Recognizing the potential for confounding or effects modification by copollutants, EPA nonetheless provisionally accepts the validity of single-pollutant PM studies when there is evidence that the effects of other pollutants are not likely to mask the PM effects. Moreover, when evaluating studies as a group as was done in the Staff Paper, similar PM associations are seen in areas with substantially varying amounts of co-pollutants, including high and low SO_2, CO, NO_2, and O_3 (U.S. EPA 1996b; Figure V-3).

55. G&L P 848, C 1, L 36-50: "When pollutants are included in the model, measurement errors ... collinearity ... may make it impossible to assess the independent effects ... in a time-series study."

RESPONSE: This is indeed a theoretical possibility. However, there is not sufficient reason to believe that measurement errors for ambient PM_{10} or ambient $PM_{2.5}$ are so large as to seriously bias the estimates of health effects from exposures to ambient PM. Studies have been carried out in many locations, with and without copollutants present, with quantitatively similar estimates of PM effect size. While it may be difficult to estimate PM effects apart from those of copollutants in some individual time series studies, there are other studies in which there appear to be distinct and separable PM effects. As discussed in #54 above, taking the entire body of epidemiology time series studies as a whole, there appears to be some adverse health effects attributable to PM or to the air pollutant mixture in many diverse locations that is well indexed by PM.

56. G&L P 838, C 2, L 8-11: "Because correlation does not prove causation in observational studies, it is necessary to evaluate these associations using Hill's criteria (1)".

This comment is repeated here because one of Hill's criteria for assessing causality was overlooked in the reviews and commentary on PM epidemiology by Gamble and Lewis and other commenters. That is, what Hill termed "experiment," through intervention that significantly changes exposure to the agent in question. The classical example involves study of associations before and after an intentional control or mitigation program removes or reduces the risk factor. Two significant kinds of examples of marked "interventions" for PM have been studied. One, discussed extensively in earlier Criteria Documents is the 1952 London episode, as well as other historic episodes, in which the "intervention" was actually a very large short-term increase in exposure over previously recorded levels. As discussed more fully above, this episode was accompanied by substantial increase in mortality and morbidity in the same population groups that appear sensitive to much lower contemporary concentrations.

A second example, more in line with a conventional intervention analysis, is provided in the recent Utah valley studies. Several time series studies in Utah valley, where PM levels were high but concentrations of gaseous pollutants are relatively low, have found associations between daily changes in PM and indices of mortality and morbidity. During the time periods studied, PM_{10} concentrations in Utah County were influenced by emissions from a large industrial source complex. The source was partially closed during a labor dispute that lasted for a 13-month period in 1986-87. During this period, PM_{10} levels were about one third lower in the community than earlier or later. Immediately before the shutdown period (April 1985 and July 1986), the level of the 24-hour standard for PM_{10} (150 $\mu g/m^3$) was exceeded 13 times; after the source reopened, 10 exceedances of the standard were recorded between September 1987 and February 1988. In the intervening months, July 1986 to August 1987, there were no exceedances of the PM_{10} standard. Studies of these periods found clear decreases in morbidity (Pope, 1989) and mortality (Pope et al., 1992) during the period of reduced emissions from the source.

The number of hospital admissions for asthma and bronchitis in children winter months when the source was closed was decreased three-fold as compared to the same season when the source was operating; the association was much stronger for children than for adults (Pope, 1989). Comparison of Utah Valley with neighboring areas not influenced by the source emissions (Salt Lake and Cache valleys) serve to confirm that other factors in the region were not responsible for the changes observed during the shutdown. Decreases in asthma or bronchitis hospital admissions found in Utah Valley were not found in these neighboring areas during the shutdown period.

In contrast, there were significantly higher per capita respiratory hospital admission rates in Utah Valley than in the Cache valley when the source was operating, even though Cache valley had a higher smoking rate and lower winter temperatures. These differences narrowed considerably when the source was closed (Pope, 1991). Moreover, Pope et al. (1992) found that the average number of daily deaths were 3.2% higher during the time periods when the source was open than when it was closed. This is quantitatively consistent with the change expected based on the observed relationship between PM_{10} concentrations and mortality across the full study period (RR = 1.16 per 100 $\mu g/m^3$ increase in PM_{10}).

This quasi-experimental study has several notable features. Mortality and morbidity rates were compared between Utah valley and the neighboring communities, and decreases in both were found only in the Utah valley population during the source closure. It has been proposed that epidemics of viral infections are a confounding factor in studies of hospitalization, but Pope (1991) found that hospital admissions were reduced during the source closure only in Utah Valley County, not in the neighboring counties. This argues against both viral epidemics and weather/seasonal changes as confounders in the relationship between PM and health, because the health benefits seen during the source closure were found only in the Utah Valley community. Pope (1991) also obtained information on economic indicators that discredit the hypothesis that economic hardship resulted in reduced access to health care, resulting in reduced hospital

admissions. Several observations argue against the economic hardship hypothesis: (1) the workers affected by the closure maintained medical insurance benefits throughout the study period; (2) only approximately 2% of the local work force was employed by the source; (3) aggregate economic indicators suggested that the impact of the closure on the local economy was minimal.

The results of these studies (Pope, 1989; Pope, 1991; Pope et al., 1992) of the intervention in Utah Valley County lend strength to the causal association between PM pollution and health effects. Overall, the studies found increases in both mortality and morbidity with increases in PM_{10} that are consistent with findings from studies in other locations. During the period of the source closure, when PM_{10} levels were reduced by approximately one-third, health benefits (reduced mortality, hospitalization, and symptoms in children) were found in the community nearest to the source but not in neighboring counties. While most epidemiological studies have consistently found decrements in health to be associated with increases in PM pollution, this intervention study presents strong evidence for improved community health when PM pollution is reduced by intervention that is unrelated to changes in meteorology or other factors that might otherwise be related to day-to-day changes in pollution. As Hill points out, these kinds of observations in intervention analyses provide some of the "strongest support for the causation hypothesis..."

57. G&L P 848, C 1: **Summary and Conclusion.** "A primary author of the PM studies concludes that the evidence seems to leave little room to doubt that particular air pollution at commonly occurring levels is causally associated with a range of adverse outcomes (13). However, . . . the causal criteria are not met and the weight of evidence does not support the PM/mortality hypothesis."

RESPONSE: Based on the far more comprehensive and incisive analyses contained in the EPA PM CD and PM Staff Paper (as reviewed by and closed on by CASAC), EPA arrived at the conclusion that the overall evidence provided by available epidemiologic studies pointed toward a likely causal relationship between ambient PM exposures and increased morbidity and mortality. The Staff Paper proposed options for the EPA Administrator to consider for revising the PM standards based on such a conclusion. Contrary to the Gamble and Lewis arguments (responded to above) EPA believes that the Hill criteria for causality are sufficiently met by the latest available epidemiologic studies to conclude that the current PM NAAQS do not provide adequate protection for public health and, therefore, require revision.

B. Review of Comments by EOP Group on Behalf of The American Petroleum Institute (IV-D-2247)

The same format as used above is employed here, with the more salient issues raised by the EOP Group presented first (by page #), followed by EPA's response.

1. EOP. P 1: "... EPA assumed a causal relationship between PM and mortality. Reliance on the Pope [1995] and Dockery [1993] studies, however, is unfounded... Because epidemiology studies often rely on imprecise data, relative risks below 2.0 require greater scrutiny. The Pope and Dockery relative risks are 1.17 and 1.26, and should be carefully scrutinized... In generally-accepted epidemiology, relative risks must be large enough to overcome deficiencies common to the epidemiologic method (e.g., greater than 4.0)"

RESPONSE: The EOP commenters have given the impression that these two studies exist in an absolute vacuum of other evidence. In fact, these are two relatively small components in an extensive historical record of evidence that has been assessed in detail in the 1996 EPA PM CD and in earlier PM CDs, addenda and supplements in 1982, and 1986. As noted earlier in response to Gamble and Lewis, abundant clinical and medical evidence was provided by particle-related air pollution episodes in the Meuse Valley of Belgium in the 1930's, in Donora, PA, in 1948, and in numerous air pollution episodes in London in the 1950's and 1960's. Table 1-12 (also Table 14-1) in the 1982 PM CD lists estimated excess mortality during 7 air pollution episodes of 4 to 6 days' duration from Nov., 1948 through Dec., 1962. The Dec., 1952 episode was so serious that there was an estimated increase of 4000 deaths during 4 days, representing an excess risk of about 3.5 or a RR of about 4.5 for each day of the episode. No sophisticated statistical analyses were needed, although a number of other investigations precluded weather or environmental stressors other than air pollution as likely causes. Medical observers suggested that elderly people with pre-existing cardiovascular or respiratory conditions were particularly likely to show symptoms of respiratory distress. Both PM levels (indexed by British Black Smoke or BS) and SO_2 levels were elevated during the episodes.

The London mortality studies cited in the 1986 PM CD provided virtually indisputable evidence that excess mortality was associated with airborne particles down to concentrations of British Black Smoke (BS) at levels as low as 150 $\mu g/m^3$. Potential confounders such as temperature, humidity, and copollutants (SO_2) were found to be incapable of explaining the excess of winter daily deaths. Initial EPA analyses were presented in the 1986 Second Addendum to the PM CD. Figures 1 and 2 (pages A-6 and A-7) show a clear concentration-dependent effect of BS on mortality in London winters 1958-1972, and this relationship persisted even after statistical adjustments for weather and SO_2. The response in population excess mortality was smaller than in earlier episodes, but quantitatively consistent with the lower BS concentrations. The excess mortality only once reached about 140 deaths per episode day during the Dec. 1962 episode, or RR about 1.5, during the 1958-1972 period. This provides a quantitative connection between earlier episodes and excess mortality at lower concentrations.

The evidence for short-term excess mortality associated with PM exposure has increased many-fold since the 1986 PM CD. Quantitatively similar findings have been found in many cities in the U.S., Canada, Europe, and Latin America, including locations where confounding from copollutants such as SO_2 is likely to be minimal. Daily time series mortality studies examine the changes in mortality rate in relation to changes in environmental stressors over periods of a few days preceding the response day, thus are not likely to be confounded by socio-demographic

factors or by changes in behavior at a population level. However, the baseline rates for daily mortality studies may include delayed responses to PM and other air pollutants, thus RR includes only the detectable top end of potential responses to PM exposure. Other epidemiology designs are useful as supplements to the findings of the short-term excess mortality studies. The studies by Dockery et al. (1993) and by Pope et al.(1995) are entirely consistent with the much larger body of evidence derived by medical observation and from short-term epidemiology studies. The EOP Group appears to be unaware of this chain of scientific evidence, which connects events with higher RR values that they would find acceptable, to qualitatively and quantitatively similar findings at much lower concentrations, even below the current NAAQS.

2. EOP. P 1: " Pope and Dockery attribute individual mortality without having measured how much individual study subjects were exposed to PM."

RESPONSE: Ambient PM_{10} and $PM_{2.5}$ have a relatively uniform regional distribution (Burton and Suh, 1995) and appear to be largely independent of PM generated indoors (Ch. 7, PM CD). Stationary air monitors provide adequate information for evaluating individual exposure to thoracic particles and respirable particles of outdoor origin. See above Appendix A responses to Gamble and Lewis on the same issue and, also, the discussion in Appendix D.

3. EOP. P 1: "Ruling out confounding risk factors and bias is key ..."

RESPONSE: The prospective cohort studies of Dockery and Pope were evaluated in detail in the PM CD and some of the EOP Group's concerns were discussed there. These studies did control for a number of subject-specific factors, particularly for individual smoking status. The findings of the prospective cohort studies are roughly consistent with those of the daily mortality and morbidity studies that are not likely to be biased by the factors described by EOP.

4. EOP. P 1: "Recent epidemiologic literature is replete with unreliable associations. ... these statistical associations have been identified, no actual cause-and-effect relationships exist."

RESPONSE: If there were only a single prospective cohort study with RR = 1.17, and no other evidence of any sort whatsoever, more caution might be justified. In fact, there are many time series studies, different prospective cohort studies with similar findings, a large number of hospital admissions, lung function, and symptom studies that establish a basis for coherence, and some recent studies that may provide a biological or mechanistic basis for these findings (see Appendix C). The real question now is not the existence of PM health effects, but rather, the estimation of the magnitude of effects occurring below the NAAQS.

The statistical analyses of the epidemiology data not only support a finding of likely causality, but also: (1) help quantify the relationship between PM and adverse health effects at levels below the current NAAQS, knowing that health effects are smaller and therefore harder to detect at lower concentrations; (2) assist in evaluating the predictiveness of PM_{10} as a gravimetric index

of thoracic particle exposure vs other alternatives, such as $PM_{2.5}$; (3) provide better characterization of adverse effects that may be experienced by susceptible subpopulations.

5. EOP. P 3: "*Black box epidemiology disparages understanding. It takes shortcuts to be able to issue 'warnings', which because of the studied 'exposures' often overlap with exhortations of politically correct moralists.*" [authors' emphasis]

RESPONSE: EPA believes that this comment is inconsistent with the facts and inappropriate in the context of a scientific review.

6. EOP. P 3-4: Bradford Hill's 'strength of association' criterion,

RESPONSE: EOP left out a critical qualifying statement that Bradford Hill included: "We must not be too ready to dismiss a cause-and-effect hypothesis merely on the grounds that the observed association appears to be slight. There are many occasions in medicine when this is in truth so." In view of the large number of potentially exposed people, EPA does not find that the observed RR are negligible. Also, as noted earlier in responses to Gamble and Lewis issues #2 and #16 above, the causal relationship of adverse human health effects due to exposure to ambient air containing high concentrations of PM was established beyond any reasonable scientific doubt by the historic air pollution episodes.

Numerous recent studies have provided evidence of causality at lower concentrations. For example, the major PM source curtailment in the Utah Valley studies discussed in #56 of the response to Gamble and Lewis above provide the results of a quasi-experimental "intervention" of the kind considered to be very important by Hill in establishing causality. Moreover, several epidemiological studies conducted at low PM concentrations in U.S. and European cities, including both short- and long-term exposures to PM air pollution, find relatively strong statistical associations with relative risks of respiratory symptom categories in children in the range of 1.5 to 5 (Schwartz et al., 1994; Pope and Dockery, 1992; Braun-Fahrlander et al., 1992; Dockery et al., 1989; Dockery et al., 1996). Concentrations in these studies extend from moderately above to well below those permitted by the current PM_{10} standards. While most of the recent epidemiological studies of mortality and hospital admissions report comparatively small relative risks, the findings of relative risks well in excess of 2 for earlier studies of high PM episodes, as well as the relative risks of 1.5 to 5 reported in more recent studies of less serious, but still important effects categories, lend credibility to EPA's interpretation of the results with respect to causality.

7. EOP. P 13: "The increased mortality claimed in the Pope and Dockery studies may have resulted from arbitrary definitions of 'most polluted' and 'least polluted' areas."

RESPONSE: See Table 12-19 in PM CD. These findings were evaluated by fitting a number of log-linear models, and retained their quantitative size and significance.

C. Review of Comments by Greenberg, R., J. Mandel, H. Pastides, L. Rudenko, and T.B. Starr on Causality, Submitted by A.P.I. (IV-D-2247)

1. Greenberg, et al. P. 1: "From an epidemiologic perspective, one principal challenge to the assertion of causality is the remarkably weak level of the reported associations. Based on the criterion of strength of association, it is difficult to imagine a weaker case for causality ..."
P 7: "This is not to say that small changes in total mortality/morbidity, if causal, are unimportant from a public health point of view. ... the issue is whether such an association is likely to be one of cause-and-effect."

RESPONSE: See responses numbered above as #2 for Gamble and Lewis, and #1 for EOP.

2. Greenberg, et al. P 1: "A second major challenge relates to the nature of PM exposure, which invariably occurs in combination with exposure to other air pollutants ... this mixture's composition varies ... extremely difficult to disentangle the potential adverse health effects of PM ... other pollutants."

RESPONSE: The most important factors in the correlation of PM concentrations are: (1) underlying meteorological conditions such as temperature inversions make concentrations of all pollutants high at the same time; (2) common emission sources for many pollutants; (3) some gaseous pollutants are precursors of PM components, such as SO_2 and sulfates. Because of the close causal connection of PM to gaseous pollutants by reasons (2) and (3), PM may be a useful surrogate for the effects of the other pollutants, and controlling the common sources of PM and other pollutants is appropriate. In view of the wide differences among communities in which significant health effects of PM have been found, it is likely that some or most adverse health effects of PM cannot be attributed to copollutants.

3. Greenberg, et al. P 1: "... most of the studies reporting significant PM associations have not included adequate adjustments for air pollution components other than PM .... (1) ... incomplete measurements on the copollutants, and (2) difficulties associated with estimation in the statistical models. ... any inference on the contribution of particulates or a specific fraction thereof is largely conjecture;" and

4. Greenberg, et al. P 9: "... the use of levels of a specific air pollutant constituent, or preferably a combination of the constituents, may serve as a very effective surrogate for overall air pollution."

RESPONSE: Many recent studies have included adjustments for copollutants. Availability of daily air pollution data for such analyses is probably even better than availability of daily PM_{10} data in U.S. cities. Statistical methods allow a substantial degree of separation of at least certain air pollutants from PM and others. For example, Samet et al. (1996b) reanalyzing Philadelphia mortality data found that O_3 and lagged CO had significant effects that were largely separate, while additive effects of TSP, SO_2, and NO_2 could not be as easily separated (SO_2 and NO_2 form

secondary sulfates and nitrates, components of PM, thus might be regarded as partial precursors of PM). Ch. 12 of the PM CD provides extensive assessment of (Samet et al., 1996b) and other studies with multiple pollutants. In some studies (such as Ostro et al., 1996), the PM_{10} effect size and significance was little changed by including copollutants in the model. In some cases, PM may be an index that includes other components of the air pollution mixture, whereas in other cases, the PM effect estimate appears to be largely independent of the effects of copollutants. The latter group of studies lend credibility to the interpretation that even when PM effects cannot be statistically separated from those of other pollutants, at least of the estimated PM effects in those studies as well are PM effects that are not included in copollutant effects. The control of PM would likely be effective in reducing adverse health effects, whether PM is a good marker for a mixture of pollutants generated by the same sources, or an independent risk factor.

5. Greenberg, et al. P 2: "... 'ecological' exposure estimates may misrepresent the association's true strength. ... the shape of the underlying causal dose-response relationship also be significantly distorted ..."

RESPONSE: See extensive responses provided elsewhere on this issue (e.g., Appendix A.II.A and Appendix D). Briefly, use of community exposure monitors for PM (especially FP) is adequate for characterizing *daily changes* in average population exposure to ambient PM, as required by daily time series analyses for changes in SMSA-wide mortality or hospital admissions rates.

6. Greenberg, et al. P 2: "Temporal relationships ... are not well established across studies ... lag times varying ..."

RESPONSE: There may be real differences among communities in terms of disease incidence in underlying populations and in the duration of weather- and pollution-related events. These will affect the timing and frequency of exposure and subsequent disease responses, thus one may expect to see real differences in lag structure among studies. Again, see extensive discussion elsewhere on this issue (e.g., Appendix A.II.A.; Appendix B).

7. Greenberg, et al. P 6: "... the conceptual approaches, study designs, and statistical analyses used across the studies are quite similar ..."

RESPONSE: Even studies by the same investigators show important differences in analytic methodology. These commenters may see similarity in approaches, but careful evaluations have identified important differences among studies that require careful evaluation, as in Sec. 12.6 of the PM CD.

8. Greenberg, et al. Pp 7-8: "... analysis of pollution data in Los Angeles County (Kinney and Ozkaynak , 1991) ... with additional adjustments for the effects of sulfur dioxide, carbon monoxide, and nitrogen dioxide, the association of mortality with particulates became non-significant statistically, with even a reversal in sign of the PM regression coefficient ... Daily

levels of particulate air pollution in Los Angeles ... were highly correlated with those for sulfur dioxide, carbon monoxide, and nitrogen dioxide (correlation coefficients for the filtered variables: 0.68, 0.82, and 0.88, respectively."

RESPONSE: The high correlations between PM and CO, PM and NO_2, suggest that a significant factor may have been the common emissions of CO, NO_2, and PM by motor vehicles. The high correlations suggests a significant collinearity among PM and copollutants that causes severe statistical instability of regression coefficients, in which a reversal of the sign of some of the air pollutant coefficients would not be unexpected. This problem can be readily diagnosed, and in some cases, substantially remedied.

9. Greenberg, et al. P 10: "... other environmental factors ... such as pollen counts and infectious disease outbreaks ... were not directly considered in these analyses. ..."

RESPONSE: Most investigators have used temporal detrending to adjust time series data for events of several weeks' duration. Nonparametric smoothers or filtering are methods that may be able to deal with transient but moderately persistent events, such as pollen seasons and epidemics.

10. Greenberg, et al. P 10: "Lipfert and Wyzga (1995a) have shown that sedentary life style ... correlated well with the mortality rate ratios observed in the Six Cities Study."

RESPONSE: This issue is responded to elsewhere (Appendix A.II.A.; Appendix B). Briefly here, life-style data were only available to Lipfert and Wyzga on a very aggregated statewide scale, thus incommensurable with regional or SMSA data. It is not clear if sedentary life-style data for each particular city in the Harvard 6-City Study would, in fact, match the state-wide average used by Lipfert and Wyzga or, if notably different, whether the apparent relationship paralleling PM concentrations would disappear.

11. Greenberg, et al. P 10: "... the persons within a community who are most heavily exposed to particulates may not be the ones who are developing the disease(s) of interest, even if more adverse health outcomes occur on days of high pollution of in areas of high pollution."

RESPONSE: PM tends to have a reasonably uniform regional distribution, and day-to-day changes in PM concentrations are reasonably adequate indicators of changes in exposure to daily ambient PM by almost all individuals in the population. Daily changes in mortality or hospital admissions rates over periods of a day or two more likely reflect changes in environmental stressors (PM, weather, copollutants) over the preceding few days. The commenters have not presented any plausible candidates for other risk factors that could have the same effects.

12. Greenberg, et al. P 11: "... community sampler measurements rarely provide good estimates of individual exposures. ... a smoker in the house provided the dominant source of indoor air pollution."

RESPONSE: Exposure to PM of ambient origin is almost completely independent of exposure to indoor-generated particles. Therefore, epidemiology studies relating changes in health endpoint to changes in ambient PM are unlikely to be confounded by the large differences among individual PM exposures due to differences in indoor-generated or personal PM sources. The key is that individual exposures to PM of ambient origin (including both while indoors and outdoors) do appear to be correlated with ambient PM concentrations monitored at outdoor sites adjacent to subjects' homes.

13. Greenberg, et al. P 13: "... reanalysis of the Philadelphia data conducted by the HEI (1996) also provided evidence of a non-linear dose-response relationship ..."

RESPONSE: The evidence on non-linearity (including Samet et al., 1995, 1996b) was reviewed in the PM CD and is not clearly resolved at this point. Evidence for a PM_{10} threshold is somewhat equivocal, with some formal hypothesis tests of nonlinearity having negative, and some positive, outcomes. The study by Schwartz et al. (1996) provides little evidence for a $PM_{2.5}$ threshold.

14. Greenberg, et al. P 20: Table 1, Studies Evaluated by Expert Panel.

RESPONSE: This is a very small list of studies, with little indication of why these particular studies were selected and numerous others cited in the PM CD were ignored (for example, the list does not include Schwartz et al., 1996, which was also reviewed). In particular, the Panel has ignored all of the studies published before 1990. EPA's CD, in contrast, represents a comprehensive review of the literature, and the studies used in the CD have undergone additional analysis and review by EPA scientists, CASAC members, and in public meetings.

D. Review of Comments by Environ Corporation (EC) on Behalf of Kennecott Copper (IV-D-2213)

1. EC. Pages 1-2: "... (1) the qualitative attribution of causality has not been clearly established ... using the widely accepted Bradford Hill criteria ... When associations between exposures and outcomes are strong, i.e., when relative risks are high, attribution of causality is relatively straightforward. When the association is weak, however, as in the case of PM, sound science requires that more emphasis be placed on making certain that the quality of the studies and their subsequent analyses meet minimal standards (i.e., reproducibility). ... The strength of the association between measures of air pollution, or any component of air pollution and adverse health outcome is very weak (relative risk ranges from approximately 1.06 to $\leq$ 1.5)." (also Sec. IIIA, P 6)

EPA RESPONSE: The commenter is aware of the extensive historical record of evidence that has been assessed in detail in the 1996 EPA PM CD (cited in his references) and in earlier PM CDs, addenda and supplements in 1982, and 1986. The commenter states on Page 2 that "...

exposures to high level of pollutants have clearly caused death and respiratory distress in the past ..." EPA agrees. More specifically, the London mortality studies cited in the 1986 PM CD provided virtually indisputable evidence that excess mortality was associated with airborne particles down to concentrations of British Black Smoke (BS) at levels as low as 150 μg/m3. Potential confounders such as temperature, humidity, and copollutants (SO_2) were found to be incapable of explaining the excess of winter daily deaths. Initial EPA analyses were presented in the 1986 Second Addendum to the PM CD. Figures 1 and 2 (pages A-6 and A-7) show a clear concentration-dependent effect of BS on mortality in London winters 1958-1972, and this relationship persisted even after statistical adjustments for weather and SO_2. The response in population excess mortality was smaller than in earlier episodes, but quantitatively consistent with the lower BS concentrations. The excess mortality from the adjusted regression models predicted about 110 to 140 deaths per episode day during the December 1962 episode, or RR about 1.4 to 1.5. This establishes a plausible quantitative connection between earlier episodes (1948-1957) and those observed during the 1958-1972 period.

The evidence for short-term excess mortality associated with PM exposure has increased many-fold since the 1986 PM CD. Quantitatively similar findings have been found in many cities in the U.S., Canada, Europe, and Latin America, including locations where confounding from copollutants such as SO_2 is likely to be minimal. Daily time series mortality studies examine the changes in mortality rate in relation to changes in environmental stressors over periods of a few days preceding the response day, thus are not likely to be confounded by socio-demographic factors or by changes in behavior at a population level. However, the baseline rates for daily mortality studies may include delayed responses to PM and other air pollutants, thus RR includes only the detectable top end of potential responses to PM exposure. Other epidemiology designs are useful as supplements to the findings of the short-term excess mortality studies. The studies by Dockery et al. (1993) and by Pope et al.(1995) are entirely consistent with the much larger body of evidence derived by medical observation and from short-term epidemiology studies.

The reproducibility of a number of studies has been independently verified by the Health Effects Institute (Samet et al., 1995, 1996a, b). Even using a different and very comprehensive model search strategy, Samet et al. (1996b) found similar effects of TSP on mortality as did Schwartz and Dockery (1992b). The HEI study largely supersedes earlier studies finding somewhat different results (Moolgavkar et al., 1995a; Li and Roth, 1996).

As other commenters addressed above, E.C. tends to ignore both Hill's cautions with respect to the strength of the associations. They also ignore other important aspects of the evidence for causality in the recent literature, notably the Utah Valley intervention studies (see #56, Gamble and Lewis above) as well as a number of recent studies that provide evidence of relatively strong statistically significant associations for some indicators of morbidity (see EOP #6 above).

2. EC. P 2: "Significant uncertainties exist in estimates of exposure. ... the ecological fallacy ... is clearly in effect because estimates of exposure in all of the studies are derived from community samplers ..." (Also Sec. IIIB).

RESPONSE: This issue has been addressed in detail in other responses above. Briefly, the use of stationary air monitors in daily time series epidemiology studies to associate *changes* in the population rate of adverse health effects on a given day to *changes* in levels of environmental stressors over the preceding few days. This can be done by statistical adjustments for weather, time trends, and air pollution. Population characteristics remain largely unchanged during such short time intervals, so no adjustment for *change* in demographic factors is needed. Recent studies (Burton et al., 1996) have shown that ambient $PM_{2.5}$ has a highly uniform spatial distribution across a large region or SMSA such as Philadelphia, and even the coarse fraction (CF) of PM_{10} (= $PM_{10-2.5}$) has a somewhat uniform spatial distribution. Ambient fine particles readily penetrate into homes, as do ambient CF particles. Thus, the use of one or a few community air samplers is adequate to characterize individual exposure to ambient particles, and in particular, to characterize *changes* in individual exposure to ambient particles over a few days on a population basis. In this sense, daily time series studies using community air samplers do not suffer the exposure characterization problem attributed to ecological studies by some authors.

3. EC. P 2: "Because $PM_{2.5}$ is largely a mixture of different chemical species arising from volatilized materials, PM precursors are not only sources, but also copollutants ... When the role of these copollutants is carefully considered, measures of PM-associated risk decrease, often to insignificance. Other pollutants ... often maintain a significant association with adverse health outcomes ..." (Also Sec. IIIA).

RESPONSE: This issue was discussed in detail in the PM CD and other responses. The discussion in Section 12.6 of the PM CD, particular relating to recent HEI findings for Philadelphia (Samet et al., 1996b), show that TSP effects can often be clearly separated from copollutants such as CO and O_3. In Philadelphia, TSP effects cannot be as readily separated from SO_2 and (to a lesser extent) NO_2, which are precursors of $PM_{2.5}$ components. However, quantitatively similar PM_{10} effects have also been found in other studies in locations where SO_2 is largely absent (Pope et al., 1992; Pope and Kalkstein, 1996), and are not attributable to O_3 in summer. In the London studies cited in the 1986 PM CD, it was the SO_2 effects that consistently were diminished while TSP effects remained strong and only slightly less significant when SO_2 was included as a copollutant than otherwise. In view of the rather small (but non-negligible) PM effects on health that are being evaluated for PM levels below the current NAAQS, including correlated copollutants in a regression model for health effects would be expected to somewhat reduce the estimated PM effect, and to increase its uncertainty, thereby reducing its statistical significance. This sometimes does occur, but in other studies, inclusion of copollutants has little effect on the size or significance of the estimated PM effect.

4. EC. P 2: "Although exposures to high levels of pollutants have clearly caused death and respiratory distress in the past, there is no compelling biological or toxic mechanism that can explain the adverse health outcomes noted for current PM levels. ..." (Also Sec. IV)

RESPONSE: Recent health and toxicology studies have identified a number of plausible hypotheses for underlying mechanisms. The diagnoses where most of the PM-related effects are now found are in respiratory and cardiovascular causes, just as in the earlier episodes. See also response to Gamble and Lewis #36 and #33 for an expanded discussion this issue, including the limitations of the controlled human and animal toxicological studies for drawing conclusions with respect to elucidating mechanisms.

E. Review of Sapphire Group Comments on Behalf of American Automobile Manufacturers Association (IV-D-2243)

1. SG. Pp 3-4: "... the Agency ... has chosen to rely on toxic potency values that fluctuate according to geographic location and population characteristics. To the best of our knowledge, this type of approach has never been utilized by the Agency for regulatory decision making ..."

RESPONSE: Site-specific factors play a major role in Superfund remediation goals for lead at CERCLA sites, as just one example. Different PM characteristics, different copollutant mixes, different climatic conditions, and populations with different age structure, baseline incidence of illness, and climatic adaptation explain at least part of the differences found among epidemiology studies. Moreover, as required by Section 108 of the Clean Air Act, the science review for all criteria pollutants has always considered the potential modification of the effects of pollutants by variable factors, including the presence of other pollutants. Commenters are clearly unfamiliar with past regulatory decisions for criteria air pollutants, and of PM in particular.

2. SG. P 4: "... the Agency's risk assessment does not identify a dose-response function ..."

RESPONSE: The OAQPS Staff Paper discusses PM concentration-response functions in great detail, based on Ch. 12 and 13 of the PM CD. While there are some uncertainties in the shape of the response functions derived from epidemiology studies, these uncertainties are described in detail and evaluated by appropriate sensitivity analyses. There is much less uncertainty than in some other kinds of risk assessment, for example, those which rely on with cross-species extrapolation.

3. SG. P 6: "... adequacy of a database ... precedents set by ... Cancer Risk Assessment Guidelines ..."

RESPONSE: The legislative requirements in dealing with Criteria Air Pollutants involve a very different set of evaluations than those involved in assessment of potential carcinogens. The commenters have at several places tried to force EPA's assessment of PM into this inappropriate framework. Extensive precedents for database adequacy for Criteria Air Pollutants have been established by many Criteria Documents over the preceding 20 years. CASAC closure on the PM CD in March, 1996, confirms that EPA's review of the scientific database is consistent with

previous NAAQS reviews and is scientifically sound and adequate as a basis for the current PM NAAQS decisions.

4. SG. Pp 7-8: "The USEPA has delineated several elements critical to assessing the hazard of a carcinogen ... *Mode of action information* ... for selecting concentration-response approaches ... *Identification of bias and adequate characterization of exposure* for assessing adequacy of epidemiological studies ... *Criteria for causality*, including biological plausibility and coherence ... *Independent toxicology studies with consistent results*, with multiple observations across species, sites, and genders. *Human clinical data which support findings* of other types of studies."

RESPONSE: PM is regulated as a Criteria Air Pollutant, not as a carcinogen. These commenters are clearly trying to force PM assessments into an inappropriate regulatory assessment framework with which they are most familiar. This framework was developed for addressing large numbers of specific toxic substances. The characteristics of the evidence for carcinogens often involves extrapolation of high dose animal experiments to humans exposed at much lower, sometimes several orders of magnitude lower concentrations. The evidence is sometimes supported by occupational epidemiological studies, but community studies are much more rare. By contrast, in examining the six criteria pollutants over the years, EPA has developed an orderly, extensive, and rigorous framework for assessing the scientific evidence and reviewing the standards. This approach has been widely accepted by the courts, CASAC, and members of the interested public. While EPA does not accept this carcinogen-based framework as appropriate for PM or other criteria pollutants, the following responds to the key points raised by this group.

Mode of action information ... for selecting concentration-response approaches .

Concentration-response functions were derived from human epidemiology studies. Low-dose extrapolation is not needed, since the range of concentrations being evaluated in the EPA proposal are well within the range of PM concentrations to which human populations in the U.S. and elsewhere have been exposed and are currently being exposed. Therefore, important issues such as the low-dose nonlinearity of the concentration response function (including thresholds) can be evaluated without use of any extrapolation model, either from human exposures at much higher concentrations, or from animal studies. A number of investigators have considered threshold models for PM, but there has so far been little empirical evidence for a PM threshold or for large deviations from linearity at PM concentrations below the NAAQS.

Identification of bias and adequate characterization of exposure for assessing adequacy of epidemiological studies

Exposure issues were evaluated in detail in Ch. 7 of the PM CD. Evidence presented there suggests that community air monitors are adequate for characterizing personal exposure to ambient PM, especially for fine particles which have a reasonably uniform spatial distribution

within a region or SMSA (Burton et al., 1996) and readily penetrate into buildings, and to a lesser extent than for coarser thoracic particles. Issues of bias and confounding were exhaustively evaluated in Ch. 12 of the PM CD. Evidence strongly supports the view that several good statistical methods are available for removing the effects of weather and long-term trends as confounding factors, such as the nonparametric smoothers used in a number of recent studies (Samet et al., 1996b). Copollutants may be potential confounders in some studies, but in some other studies the PM effects were substantially separated from those of other pollutants such as CO and O_3. While it may be difficult to separate PM from some precursors such as SO_2 or NO_2 in some studies where their concentrations are high enough to cause health effects, there are other places where levels are sufficiently low that they cannot be confounding factors. Even when PM effects are partially confounded, the typical result is a small reduction in the estimated PM effect.

Criteria for causality, including biological plausibility and coherence

See several responses, including the entirety of Gamble and Lewis, #1 and #6 EOP.

Independent toxicology studies with consistent results, with multiple observations across species, sites, and genders.

Toxicology studies are reviewed in Ch. 11 of the PM CD. See Gamble and Lewis response #33 and 36 for additional discussion.

Human clinical data which support findings of other types of studies.

Human clinical data do not show large PM effects in many studies. This may represent the fact that many such studies involved healthy subjects or otherwise healthy young asthmatics. However, PM effects seem to be mainly detectable epidemiologically in elderly people, especially those with compromised health status. Clinical studies are not carried out with such subjects who may face a high risk of respiratory distress or even death from the study. Also, the brief exposure (1-4 hrs) typically used in the chamber studies may not be sufficiently long to mimic effective ambient PM exposures (1-3 days) most clearly associated with the mortality/ morbidity effects observed epidemiologically in the general population. Again, see further discussion in Gamble and Lewis response #33 and 36 above.

5. SG. Pp 8-10: "... time series studies ... deal with death *counts* rather than death *rates* ..."

RESPONSE: This issue has not been discussed in much detail because of the consensus that the time series data were appropriately analyzed. The commenters' recommendations cannot be implemented because of the problems in determining fluctuations in the size of the exposed population. Most investigators have resolved this by restricting counts to people who both lived in the region, and died or were admitted to hospital in the region. The residence information was available from NCHS data tapes in the mortality studies. Investigators have generally used

flexible methods for detrending the time series, which may implicitly capture some of the population changes over short time intervals by use of a surrogate. In any case, the effect is expected to be very small. The commenters' speculations about factors associated with the size of the exposed population are not supported by factual evidence. Most investigators now assume that daily death or hospital admissions counts follow a Poisson or hyper-Poisson distribution. The assumption of a hyper-Poisson distribution with an over-dispersion index allows further possibility of adjusting the count data for other sources of variation, such as population fluctuation.

6. SG. Pp 10-11: "*The Statistical Significance of the Time Series Studies is Difficult to Judge Because Time Series Studies Offer Too Much Flexibility for 'Finding' Associations*"

RESPONSE: The problem of model specification searches was discussed Ch. 12 of the PM CD. Most investigators now use highly directed model search strategies that greatly reduce the number of models to be evaluated (Samet et al., 1996b). Almost all of the models that best fit the data have a rather similar structure, not much different that found for the London data (Schwartz and Marcus, 1986, 1990). Some of the parameters may differ, but differences in PM lag models are not great. These differences may well reflect real differences among air pollutants and population in different cities.

F. Review of Comments by Cox Associates on Behalf of the Engine Manufacturers Association (IV-D-2328)

Key issues raised in the Cox Associates (CA) comments are numbered and identified first (by Page #), followed by the EPA response to each.

1. CA. P 1: "It would be necessary that PM_{10} and O_3 actually cause adverse health effects below the concentration levels already required by existing standards ... Allowing professional statisticians and experts in causal analysis to reanalyze these data could help to resolve the key issue of whether the statistical associations that EPA cites between PM and O_3 and adverse health effects represent true causal relationships ..." [author's emphasis]

RESPONSE: The issue of causality for PM is discussed in the full response to Gamble and Lewis above. Note especially #3.

2. CA, P 3: "... relies disproportionately upon the opinions expressed by a small but prolific community of researchers ... associations that other investigators have been unable independently in other locations (e.g. Bacharova, 1996 ... or have been disconfirmed in the same locations or using the same data (Styer et al, 1995; Davis et al., 1996) ..." repeated at p. 25 and numerous other places.

RESPONSE: EPA carefully reviewed all of the studies available at the time, and noted both strength and weaknesses of all of these studies, including detailed and objective assessments of the studies prepared by the "small but prolific community of researchers." Cox Associates fail to note that many of the findings were verified numerically by completely independent reassessments carried out by the Health Effects Institute (Samet et al., 1995) and also by independent reanalyses using different methods for Philadelphia (Samet et al., 1996a,b). In addition, Pope and Kalkstein (1996) reevaluated the Utah Valley data using alternative controls for effects of weather, including the synoptic categories proposed by Kalkstein (who is not a member of the "prolific group" referred to by Cox). The Pope and Kalkstein (1996) analyses still found significant PM effects, similar to those originally reported.

As explained in the promulgation notice, studies published after the PM CD closed were not considered in the final decision. EPA notes, however, that Cox Associates has very selectively cited European findings of the APHEA investigators such as Bacharova (1996) and of similar studies such as Ballester (1996). Most APHEA investigators found positive and statistically significant health PM effects for European cities, such as Anderson et al. (1996) for London mortality. A recent meta-analysis (Katsouyanni et al., 1997, in press) reports strong positive findings for Western European studies and weaker, but still positive, findings for central and Eastern European APHEA studies. See Appendix C for a more complete and provisional description of such APHEA results.

The studies by Styer et al. (1995) were briefly reviewed in the PM CD and found to have serious methodological limitations. The Chicago analyses by Ito et al. (1995) were assessed by EPA as more credible than the Styer analyses, but, even so, Styer et al. still found significant associations between PM and mortality when analyzed for the entire year. The methodological problems with the Styer analyses means that it did not actually disconfirm findings by Pope et al., apart from the flaws in the Styer methodology. The recent unpublished report by Davis et al. was not considered in the CD. As discussed in Appendix C, it contains several factual misstatements, and recent reanalyses have shown that use of the Schwartz PM metric produces very similar findings of a significant PM_{10} effect, even with different weather model specifications (R.L. Smith, personal communication, Feb. 8, 1997; NISS comments, March 14, 1997). See above EPA responses to G&L for refutation of Table 2 findings on CA P 25.

3. CA. P 6: "... we believe that EPA has not been sufficiently critical in deciding whether appropriate statistical methods and models have been used ..."

RESPONSE: Over 300 pages of the PM CD was devoted to assessing the epidemiology results, and the methods and models on which they were based. These issues were introduced and discussed early in the PM NAAQS review process at a PM Mortality Workshop sponsored by EPA in November, 1994. Also, CASAC reviewed the EPA PM CD methodology assessments in 1995 and 1996, and voted closure on the document. EPA has continued its methodology assessments, even on unpublished reports received as part of Public Comments on the Proposed Decision. EPA has addressed specific issues raised by CASAC and by public commenters.

Almost all of the generic issues raised by Cox were discussed in the PM CD, or in EPA responses to specific studies alluded to by commenters. Cox has not identified specific studies in which the other issues are alleged to have arisen. Numerous opportunities existed for the scientific community to comment on the studies when they were originally published and reviewed by peer reviewers, during the numerous opportunities for public comment on the studies and when they were discussed in public CASAC meetings. EPA believes that the studies cited as the basis for the final decision on the standard have received far more detailed assessment than most other scientific documents and constitute a credible basis for regulatory decisions.

G. Review of Comments Submitted by American Industrial Health Council (AIHC) (IV-D-2340)

Comment. This report updates the review paper by J.F. Gamble and R.J. Lewis (*Env. Hlth. Persp. 104:* 838-850, 1996; denoted G&L). Most of these issues were addressed earlier in the in EPA response to G&L. Among the key issues posed by AIHC are the following three:

(1) The epidemiology studies used by EPA all follow an "ecologic epidemiologic study design" without personal exposure indices, and are therefore do not provide an adequate basis for hazard identification in EPA regulatory proposals;

(2) The causal nature between PM and mortality has not been completely assessed by EPA; and

(3) Ecologic epidemiology studies should not be used for dose-response estimation in risk assessment.

EPA's responses to each these major issues are developed in the full response to G&L above. In addition, EPA offers the following points:

(1) The daily time series studies used by EPA are considerably more informative than general ecologic studies. Textbook citations ignore some important aspects of PM epidemiology. PM indices such as PM_{10} and $PM_{2.5}$ tend to have spatially uniform regional distributions, and also can effectively infiltrate most buildings. Therefore, PM measured by one or a few stationary air monitors does adequately characterize most individual exposures to ambient PM from outside sources. More importantly, PM adequately characterizes *changes* in individual daily exposure to ambient PM. Since indoor-generated PM is almost completely independent of ambient PM, the epidemiology studies cited in the proposals provide considerably more information about ambient PM exposure than most so-called "ecologic" studies.

(2) The causal nature of the relationship between PM and excess mortality and morbidity from cardiovascular and respiratory causes was established by the pollution episodes in Belgium in 1930, Donora in 1948, and London in 1952. There were many similarities between the 1952

London episode and later London episodes where PM exposure was lower, including the Dec. 1962 episode. The quantification of the relationship between mortality and PM was based on data from 14 London winters from 1958 through 1972, discussed in the 1986 PM CD. EPA showed that the clearly identifiable effects of high pollution episodes on excess mortality were quantitatively consistent with smaller PM effects at lower concentrations, down to concentrations of Black Smoke (BS) as low as 150 $\mu g/m^3$ and possibly even lower. The more recent studies reviewed in the PM CD use more appropriate PM indices such as PM_{10} and $PM_{2.5}$, and a number found relatively strong associations (RR>1.5 to 2.0) for some effects categories. In addition, the overall consistency and coherence of the over 80 studies provides relatively strong evidence for effects at levels at or below those allowed by the current standards.

(3) Much of the risk assessment paradigm advanced by AIHC is not relevant to the acute effects of short-term exposures to noncarcinogens (although some PM components may be carcinogenic or mutagenic, and long-term effects cannot be precluded). Consequently, the daily time series studies and the prospective cohort studies can be used to improve estimation of dose-response relationships in susceptible human populations, since the "hazard identification" has long since been established. See also response #4 to the Sapphire Group above.

More specific additional responses to particular points raised by the AIHC are provided below.

AIHC. P 3: "... many of the confounders (e.g. weather and co-pollutants) are measured with the same degree of error ... [as] PM".

RESPONSE: EPA has extensively evaluated the potential for measurement errors in instrumentation or exposure characterization errors to distort relationships between PM and adverse health effects. Although theoretically possible, the measurement errors do not appear to be large enough, nor do potential confounders and their measurement errors appear to be sufficiently correlated with PM and its measurement errors, to be capable of quantitatively accounting for the observed PM effects.

AIHC. P 3: "... statistical adjustment procedures cannot totally control for a confounder's effect ..."

RESPONSE: Careful analyses can come close. In a recent reanalysis of the Philadelphia TSP data, Samet et al. (1996b) showed that careful adjustments for weather, for temporal effects, and for co-pollutants, could reduce a model goodness-of-fit index to the point where it was not possible to distinguish the remaining uncertainty from pure randomness. It would be reasonable to conclude that the most important time-variable environmental stressors that are likely to be associated with respiratory and cardiovascular endpoints during intervals of a few days are included in the statistical adjustments. Although PM effects in recent studies are small, there is an extensive record of higher PM concentrations being associated with adverse health effects, so that PM is a reasonable candidate for evaluation at current levels.

AIHC. P 5: "... the results found in the PM literature seem to be most influenced by the analysis methods and interpretations of study investigators..."

RESPONSE: The recent HEI reanalyses of the Philadelphia data suggest that, given a reasonable approach for model specification search, rather similar results can be obtained by other investigators than the original investigators. However, some of the published analyses selected quite unreasonable approaches, or in some cases, no obviously cogent approach.

AIHC. P 6: "A number of facts also argue against coherence ... chamber studies conducted at much higher PM concentrations ... Shift studies conducted on small sized particles ..."

RESPONSE: Many observers of the London episodes have identified respiratory distress and excess mortality with individuals who already had compromised respiratory function. Undiagnosed cardiovascular disease may also have been present. Most clinical studies do not use individuals who are at risk of seriously adverse response. While otherwise healthy asthmatics are sometimes exposed to PM or other air pollutants, there is no reason to believe that this group is especially susceptible to PM exposure. See also response to Gamble and Lewis #33.

AIHC. P 7: "... some believe that this [temporality] is the only factor which must be established in assessing causality..."

RESPONSE: Almost all studies of temporality have established that excess mortality or hospital admissions are associated with PM exposure in the preceding few days; rarely, on the same day; and never on future days.

AIHC. P 7: "Strength of the association ... nowhere does EPA state that this is a limitation when evaluating causality."

RESPONSE: See earlier comments above responding to Gamble and Lewis.

AIHC. P 8: "... recent European studies ..."

RESPONSE: Reviews of recent APHEA studies find many positive and statistically significant relationships, especially in Western European studies as discussed below in Appendix C (Section II). Two recently published meta-analyses have found overall significant relationships.

AIHC. P 9: "... London, England in 1952 ... (45,000 $\mu g/m^3$) ... It is very likely that the London incident, and possibly others were due to high concentrations of acid aerosols ..."

RESPONSE: In fact, the maximum London BS level was about 4500 $\mu g/m^3$, not much greater than in several other episodes, and not 45,000 $\mu g/m^3$ as stated by AIHC. The acid aerosol hypothesis was evaluated in detail in Sec. 12.5 of the PM CD. The effects of BS, SO_2, and

H_2SO_4 were discussed by Thurston, Ito, and Lippman (1989) who concluded that because of collinearity during the 1963-1972 period, "... once total mortality was prefiltered for slow-moving autocorrelations, the acid relationship appeared to be roughly similar in strength to that for BS and SO_2..." The evidence for the role of acidity is still somewhat uncertain. While some studies find significant acid aerosol health effects, and these may in fact occur in some studies, PM effects are also found in studies with little acid aerosol. Acidity is a plausible mechanism, but not the only plausible mechanism.

AIHC. P 10: "Hertz-Picciotto (1995) ... epidemiologic studies ... which can be used in dose-response assessments are required to have exposure linked to individuals."

RESPONSE: PM_{10} and $PM_{2.5}$ tend to have reasonably uniform spatial distributions; in such cases, individual exposure (and more importantly, changes in individual exposure and changes in population exposure to ambient PM from outside sources), are adequately characterized by stationary air monitors.

The Hertz-Picciotto criteria cited by AIHC only refer to epidemiology data used as a basis for extrapolation (Category 1). EPA's use of epidemiology for the PM rulemaking goes beyond both their Category 1 and their Category 2 criteria: "Can be used to check plausibility of an animal-based risk assessment" alluded to by AIHC. The present EPA PM risk assessments are based on both with more recent epidemiologic data , but are quantitatively and qualitatively consistent with earlier medical direct observations and quantitative extension of direct observations from higher levels in the London studies. As such, the direct human observations used here by EPA are not subject to the considerable inherent uncertainties of risk assessments that mainly rely on animal data. As Hertz-Picciotto says about the advantages of epidemiology data, the magnitude of error is likely to be greater when animal data are used.

APPENDIX B. RESPONSES TO COMMENTS ON USE OF SPECIFIC EPIDEMIOLOGICAL STUDIES

In contrast to the more general comments on the use of epidemiology addressed in Appendix A, another group of commenters made more specific challenges to EPA's assessment of the epidemiological studies. These comments, although overlapping some of those made by the first group, were generally made by industry groups and sponsored scientists who took a more active role in the review of the Criteria Document and Staff Paper. Examples include submissions and some attachments provided by the American Iron and Steel Institute, the American Petroleum Institute (API), the Electric Power Research Institute (EPRI), and the National Mining Association (NMA).

These commenters asserted that the epidemiological evidence on PM is not as consistent and coherent as EPA has claimed, and, in particular, charged that EPA ignored or downplayed a number of studies that the commenters argue contradict the evidence the Agency cited as supporting the consistency and coherence of PM effects. The studies, all of which commenters contend do a better job of addressing one or more key issues, such as confounding pollutants, weather, exposure misclassification, and model specification, than earlier studies, include several that were available during preparation of the Criteria Document, and a number that appeared after the Criteria Document and Staff Paper were completed. Because the status of the later studies differs from that of the earlier ones for purposes of decisions under section 109, studies in the first category are discussed in section I of this Appendix, and those in the second are discussed in Appendix C. In addition to the inclusion of specific studies, commenters also raised issues regarding the limitations of the available epidemiological information. A representative set of comments, based in large part on those provided by Dr. Ron Wyzga of EPRI, is discussed in Section II of this Appendix. One subset of these issues, regarding exposure misclassification, the role of indoor air pollution, the relationship of ambient air pollution to personal exposure and measurement error, is broken out for expanded discussion and analysis in Appendix D. Finally, in response to many commenters who cited varying lists or numerical counts of "positive" and "negative" epidemiological studies on PM, Section III of this Appendix identifies and classifies 87 key community epidemiological studies listed in selected summary tables of Chapter 12 of the CD.

I. STUDIES INCLUDED IN THE CRITERIA REVIEW THAT COMMENTERS SAY EPA IGNORED OR DOWNPLAYED

Comment: Several commenters stated that EPA ignored or downplayed a number of so called "negative" studies that call into question EPA's claim of a consistent and coherent data set for PM. Key "negative" studies highlighted by various commenters include: (a) several available at the time of the PM CD criteria review, e.g., Abbey et al. (1991) for California; those of Styer et al. (1995) for Salt Lake City; Moolgavkar et al. (1995a) for Philadelphia; Moolgavkar (1995b) for Steubenville, Ohio; and (b) others only becoming available since the PM CD closed, e.g.,

Morris (1996) for Manchester, England; Burnett et al. (1997) for ten Canadian cities; Davis et al. (1996) for Birmingham, AL: Roth and Li (1997) for Birmingham, AL; Bacharova et al (1996) for Bratislava, the Slovak Republic; Roth et al. (1997) for Prague, the Czech Republic. EPA responses to comments regarding studies available at the time of the PM criteria review are provided in this section below. For reasons discussed in the preamble to the final regulation, these latter more recent studies were not considered in the promulgation decision. A provisional examination of their results is included in Appendix C below.

RESPONSE: Negative[1] as well as positive studies available to EPA at the time of CASAC clearance of the PM CD in March, 1996 were examined in detail in the PM CD. In particular, the negative findings of Moolgavkar et al. (1995ab) of the Philadelphia and Steubenville studies were discussed in great detail in Sec. 12.6 of the PM CD in comparison to those of the original investigators (Schwartz and Dockery, 1992ab) and other investigators. Further analytical studies of the Philadelphia data set (extended to more years) were carried out by HEI (Samet et al., 1995; 1996a,b; 1997) and have largely resolved many of the uncertainties in the earlier analyses. Even though TSP is not the best PM indicator for health effects, since it includes a substantial fraction of non-thoracic particles, the extended Criteria Document assessment (U.S. EPA, 1996a, pp. 12-291 to -299; 12-327) of the Phase I.B HEI analyses in Philadelphia (Samet et al., 1996a,b) serves to support the following inferences: (1) the mortality effects estimates for TSP do not depend heavily on statistical methods when appropriate models are used; (2) estimated PM effects are not highly sensitive to appropriate methods for adjusting for time trends and for weather; (3) air pollution has significant health effects above and beyond those of weather; (4) copollutants such as ozone, CO, and NO_2 may be important predictors of mortality, but their effects can be substantially separated from those of TSP and SO_2; and (5) the health effects of TSP in Philadelphia cannot be completely separated from SO_2, which is itself a precursor of fine particles, based solely on the epidemiological analyses in this single city. These recent findings tend to support many of the positive and "negative" findings of the earlier studies. In reaching decisions on the standards, EPA has relied on the qualitative findings from the entire body of

---

[1]NOTE: The term "negative" studies should not be construed to mean those in which there is either a significant or non-significant negative effects estimate ro the nominal cause. As used by most commenters, it also includes statistically non-significant positive effect estimates. Terminology used by commenters defines "positive" studies as those in which the effect estimate is both positive and statistically significant. If there were in reality no effect from PM, then one could expect about the same number of statistically significant positive effect estimates. In reality, statistically significant negative estimates for PM effects have almost never been reported in methodologically adequate analyses. Looking across all of the published studies reviewed by EPA and CASAC (as well as those published more recently), the overwhelming majority of so-called "negative" studies find positive PM effects, even if they are not statistically significant at
the 95% confidence level.

evidence, but has placed greater reliance on studies in which PM_{10}, $PM_{2.5}$, or other gravimetric indicators of fine particle mass and components.

In addition to relying on the most comprehensive and best analyses in evaluating the reanalysis in Philadelphia and other areas, the Criteria Document gave less weight to both so-called "negative" and "positive" studies with methodogical limitations. In particular, EPA agreed with the epidemiological experts on CASAC (Lippmann et al., 1996; Samet, 1995) that the Li and Roth (1995) study approach of using a "panoply" of different modeling strategies to produce seemingly conflicting findings provides little useful insight and is superseded by the HEI report. The attempt by Lipfert and Wyzga (1995) to address relative effects of different pollutants was considered inconclusive (Lippmann et al., 1996) and flawed by the use of a metric (elasticity) that gives insufficient consideration to the absolute concentrations of the pollutants being compared (see further discussion of this issue in section II of this appendix).

Further, the Steubenville studies and reanalyses (Schwartz and Dockery, 1992b; Moolgavkar, 1995b) were discussed in detail to examine methodologies, and the differences in relative risks between the two were regarded as small U.S. EPA, 1996a, p. 12-280 to 283). Both studies used TSP as the PM indicator variable, and they are augmented by the more recent findings of Schwartz et al. (1996) that examine PM_{10} and its components. The mixed results by Lyon et al. (1995) in Utah Valley are compromised by loss of information related to the methodology (U.S. EPA, 1996a, p. 12-58). As noted in Appendix A, subsequent reanalyses of the Utah Valley study by HEI (Samet et al., 1995) as well as by Pope and Kalkstein (1996) confirmed the original findings of Pope et al. (1992) using different model specifications. The Salt Lake City study by Styer et al. (1995) was mentioned in the PM Criteria Document, but received little discussion because aspects of the methodological approach limited its statistical power to detect effects. The analysis of Chicago mortality data in the same paper shared these problems, particularly for seasonal analyses; in this larger city, they nonetheless found significant associations on an annual basis between PM_{10} and mortality that are consistent with other studies. In short, the record shows that EPA did not ignore these short-term exposure studies cited by commenters; moreover, the relative weight given to them is in full accord with the written advice provided by four CASAC panel members with extensive involvement in conducting population studies of air pollution (Lippmann et al., 1996).

Similarly, EPA believes that appropriate treatment and weight were given to studies of long-term exposure and mortality. EPA concluded that the lack of associations in the Abbey et al. (1991) prospective cohort study were not inconsistent with two other such studies because: (1) the use of days of peak TSP levels as the PM indicator (instead of PM_{10} or $PM_{2.5}$) is inappropriate for California cohorts exposed to both urban smog and fugitive dust episodes; and (2) the overall sample size may have been too small to detect significant effects (U.S. EPA, 1996b; pp. V-17 to -18). The inadequacy of Lipfert's (1995) application of state-wide average sedentary lifestyle data to adjust mortality for the six cities studied by Dockery et al. (1993), in which superior subject-specific body mass index data had already been considered, was also noted and addressed in the Staff Paper (U.S. EPA, 1996b; p. V-16). Again, EPA did not ignore these

studies; the rationale for giving them less weight was clearly articulated in the documents reviewed by CASAC and judged appropriate for use in standard setting.

EPA used all of the available studies in its review of PM health effects, including the most recent studies available at the time, with more definite publication status than some of those cited by commenters. "Negative" studies were evaluated in detail along with "positive" studies when they were found to have no critical methodolological deficiencies. Studies that had more serious problems in terms of data base or methodology were discussed in less detail, whether their findings were positive or negative, than studies with fewer or smaller deficiencies in the opinion of reviewers. The EPA assessments of the studies were evaluated by peer reviewers, by CASAC, and by the public from Jan., 1995 through CASAC closure on the PM CD in March 1996. Commenters' assertion that EPA has ignored negative studies is not supported by this record.

Comment: A number of commenters also asserted that in every case where the data from an initial study have been re-analyzed, the results of the re-analysis have not unambiguously implicated particulate matter despite the fact that the initial study implicated particulate matter.

RESPONSE: EPA has extensively reviewed all of the original analyses, re-analyses, and counter-analyses provided to the Agency. No original analysis was accepted without critical review, and those for which subsequent reanalyses had been done were given additional attention so as to better evaluate any differences reported in findings among the different investigators. Because of the large differences among investigators for some studies, it is more appropriate to evaluate these differences in detail (as was done in the PM CD) for each study and city. Nevertheless, it should be noted here that certain of the more credible reanalyses (e.g., the HEI ones by Samet et al.) have, in fact, also implicated PM. Other reanalyses (e.g., by Davis, et al.; see Appendix C) initially reported as not substantiating original findings, upon further inquiry are found to have obtained similar results after all, when the same model specifications are used as by the original investigators.

Comment: Ron Wyzga noted that the Health Effects Institute study of daily air pollution and mortality in Philadelphia, after application of several models, concluded 'We caution against using the model coefficients directly to estimate the potential consequences of lowering concentrations of the individual pollutants through regulatory measures ... pollutants are correlated."

RESPONSE: The results of the HEI study, reported on by Samet et al. (1995, 1996ab) were summarized in detail in and extensively evaluated in the PM CD. As outlined above, there EPA found that TSP and SO_2 in Philadelphia are indeed highly correlated, and that their joint effects (and to a lesser extent, those of NOx) may be hard to disentangle *in Philadelphia for the study years*. The commenters did not discuss the HEI findings that effects of ozone and CO can be substantially separated from those of particles and their gaseous precursors, SO_2 and NO_x, in

Philadelphia. These studies were not included in quantitative comparisons of relative risk and related results in the summary Tables in Chapter 13 of the Criteria Document because the particle index, TSP, is not being considered as a basis for setting standards, and estimates of PM_{10} and $PM_{2.5}$ effects on mortality are available from other studies. Some PM_{10} studies have been done in locations where the levels of certain potentially confounding copollutants are so low as to be unlikely to have any substantial effect on health. The Philadelphia TSP data have proven to be most useful for comparing analytical methodologies and examining confounding and effects modification by co-pollutants.

Comment: Many commenters alluded to reanalysis by Moolgavkar et al. (1995a) and Wyzga and Lipfert (1995ab) as not substantiating originally reported findings for Philadelphia.

RESPONSE: These studies were discussed in detail in Sec. 12.6 of the PM CD. TSP studies were used to compare different methodologies. TSP studies did not play an essential role in the regulatory proposals because TSP (which includes a large mass of very coarse non-thoracic or non-respirable particles) is known to be an intrinsically less accurate index of potential adverse human health effects from inhalation of thoracic particles than PM_{10}. The HEI reanalyses (Samet, et al., 1996a,b; EPA 1996a) substantiating positive PM effects in Philadelphia are viewed by the Agency and CASAC as superseding the earlier Moolgavkar (1995a) and Wyzga and Lipfert (1995a,b) analyses.

Comment: Commenters also refer to Moolgavkar (1995b) as not substantiating originally reported findings for PM effects in Steubenville.

RESPONSE: The Steubenville TSP studies were extensively evaluated in Sec. 12.6 of the PM CD, but did not play an essential role in the PM proposals. The issue with respect to co-pollutants is more appropriately addressed in the Philadelphia data. In any case, more recent findings for PM_{10} and its components are reported in (Schwartz et al., 1996) and supersede earlier findings based on TSP.

Comment: Ron Wyzga pointed out that an additional [Philadelphia] study by Cifuentes and Lave cited by EPA was never accepted for publication and is not in press."

RESPONSE: At the time the PM CD and Staff Paper were completed, it was EPA's understanding that the paper had been accepted for publication. The paper was reviewed by CASAC and members of the public in the review, and the authors presented their results at an open CASAC meeting. EPA later learned that the paper had not been accepted. The correct current citation status should be changed to Ph.D. dissertation, Carnegie-Mellon University. This paper was in no way central to EPA's conclusions on the consistency and coherence of the epidemiological data, or on the decisions regarding the need to revise the standard or select particular levels.

Comment: Dr. Ron Wyzga of EPRI referred to a re-analysis of the relationship between daily hospital admissions for pneumonia and air pollution in Detroit originally analyzed and published by Schwartz (1994a).

RESPONSE: The reanalysis alluded to has not yet been provided to EPA or peer reviewers, so limited technical response to these comments is possible. In any case, the Schwartz Detroit study did not play an essential role in EPA PM proposals, but did play a contributory or supportive role suggesting the plausibity of pneumonia as another health effect identified with PM exposure. Day-of-the-week PM effects have been evaluated in a number of studies and are plausible, but the findings are mixed. In any case, weekend and day-of-the-week effects may be modifying factors of more or less importance, but almost certainly are surrogates for causative factors such as changes in motor vehicle or power plant emissions or time spent in indoor or outdoor microenvironments (e.g., time in motor vehicles during commuting) that reflect weekend-to-work-week transitions.

Comment: In reaching conclusions on the reanalyses, Wyzga stated that "... the reanalyses demonstrate that it is possible to obtain very different results with an equally legitimate analysis. ... particularly disturbing for those data sets that cannot be reconstructed from publicly available data. ... extreme caution must be applied to study conclusions ..."

RESPONSE: EPA has been presented with a plethora of epidemiology studies with diverse findings: usually that there is a positive and significant PM effect, but also, at times, that there is no significant PM effect, or that the PM effects cannot be separated from those of other pollutants or confounding factors. EPA reviewed these studies individually and in great detail, since it became obvious that not all of the analyses presented to EPA were equally legitimate, valid, or credible. EPA assessments of these diverse findings were laid out in considerable detail in the PM CD and again in the Staff Paper. CASAC reviewed the EPA assessments on numerous occasions, with public comments as part of their review process. A number of positive studies *and* so called negative studies were found to be credible, but the weight of evidence overwhelmingly pointed to small but detectable adverse human health effects at levels below the current NAAQS.

EPA's responses to comments on its use of published studies for which all underlying data are not readily available are discussed at length in the preamble to the final regulation. Nevertheless, it is suprising that Dr. Wyzga, who made several assertions in his submitted comments based wholly on the results of analyses and studies that, to EPA's knowledge, do not even exist as unpublished manuscripts, much less as peer reviewed papers, would caution against the use of such peer reviewed, published results.

Issues in analysis and methodology were extensively reviewed by CASAC at every stage, and most investigators now use a small number of methods and approaches that are generally agreed to be the most informative and appropriate for these data. In the process itself, extreme caution was applied to study conclusions at every stage of review. EPA believes that within the range of

legitimate strategies and methods accepted by most statisticians and scientists, there is general concordance of findings. There are still some difficult questions that require better resolution, but like any other scientific issue, not all methods will turn out to be equally valid for answering these questions, and scientific judgment will likely converge to the best methods.

II. DETAILED RESPONSES TO SPECIFIC COMMENTS ON EPIDEMIOLOGICAL STUDIES

Comment: A number of commenters referred to issues and uncertainties in the recent time-series statistical epidemiology studies. A representative comment, by Ron Wyzga of EPRI, states that there is sufficient flexibility for analyses of large data sets that it may be possible to obtain almost any result desired through choice of statistical method. In time series analyses which relate daily health events to pollution levels, these options involve choices regarding:

(a) Statistical model, e.g. use of non-parametric smoothing techniques;
(b) Methods used to adjust for seasonal variation and the trends in the data;
(c) Other pollutant variables;
(d) Other variables (e.g. weather, day of week);
(e) Lags in the pollution and weather variables;
(f) Population (age group, cause of death, etc).

RESPONSE: More specifically, commenters suggested that one or more choices in the above areas have been shown to influence the results of the PM time-series analyses. They allude, in particular, to results of reanalyses of PM data reported by: Wyzga and Lipfert (1995 a,b) as illustrating impacts of alternative choices regarding (a) and (b) above; Davis, et al. (1996) regarding (a) and (d) above; Samet, et al. (1996) regarding (c) above; and Lipfert and Wyzga (1995a, b) regarding (e); and these and others regarding (f). Commenters argued that you can get any answer you want by an equally legitimate analysis. EPA disagrees with such comments. Each of these studies have been evaluated in detail and finds problems with each. For example, the Lipfert and Wyzga studies used an inappropriate index of relationship, elasticity, as discussed below. The unpublished findings reported in Davis et al. (1996) have subsequently been modified by analyses more consistent with the methods EPA finds more appropriate (see appendix C). The six criteria proposed by Dr. Wyzga of EPRI are worthy of consideration in evaluating individual studies, but some of the investigations he cites have not used the most appropriate or powerful methods.

EPA's detailed reviews of individual studies have shown that not all methods are equally valid or legitimate. Preferred methods for dealing with each of the above six points include:

a. Statistical model

Most analysts recognize that discrete count data with a binary outcome (such as dead/not dead today) should be roughly binomial in a finite population, but that the Poisson distribution is an excellent approximation in SMSA studies. The Poisson mean depends on demographic and environmental covariates, however. The normal distribution is at best an approximation when the daily number of events is large, such as in London in the 1960's with about 280 deaths per day in baseline winter conditions. Furthermore, the Poisson distribution requires variance = mean (variance > mean for hyper-Poisson counts), a restriction which must be superimposed on normal or lognormal models. Poisson or hyper-Poisson time series methods are preferred.

b. The methods used to adjust for seasonal variation and the trends in the data

Classical (normal or log-normal) time series methods such as filtering (for example, subtracting 15-day moving averages) may be less appropriate than subtraction of long-term trends by non-parametric smoothing functions of time, with time span of about one to three months. Use of dummy variables for seasons may be adequate, but often invites over-parameterization by including interactions (i.e., seasonal differences or effects) on air pollution and weather parameter interactions. Simple detrending by linear or quadratic functions over several years may work, but also may be inadequate to capture changes in demographics or pollution mixture over time. Fourier series may assume more regularity by calendar date than actually occurs. Non-parametric smoothers seem to work best, with less potential for confounding with episodes of a few days that are governed by weather system (synoptic category) changes. Under some conditions, the direct use of the synoptic climatologic categories is also reasonable.

c. Other pollutant variables

Some investigators such as HEI have explored this issue in detail using two-dimensional nonparametric smoothing surfaces (Samet et al., 1995). The problem may be more complex than it looks, since this includes using air pollution metrics with the appropriate lag or moving average structure, see E below, and pollution effects may change by season. For PM, the London studies show that there are significant effects from BS (PM) exposure on the two previous days and a smaller effect for current-day exposure, whereas ozone effects seem to described best by same-day or previous-day concentrations. Assuming the correct lagged metric is used, classical chemometric methods for decomposing a multivariate lognormal distribution of correlated air pollution concentrations could be used to better assess which pollutant combinations are profoundly correlated with others and which are not. The standard method for assessing confounding in the air pollution health studies reviewed by EPA is to examine the sensitivity of the PM regression coefficient to inclusion of other pollutants. This is an appropriate method to examine copollutant effects when the health response of interest (such as cardiovascular or respiratory mortality in the elderly) has a reasonably log-linear relationship to the pollutants, but is not such a good idea with a profoundly nonlinear relationship (such as a "threshold" model). Goodness-of-fit assessments against nonparametric smooth functions of air

pollution are feasible, but rarely done. The few analysts who have pursued these model specification tests tend to find that a linear or piecewise linear model for PM_{10} is often adequate (i.e., not providing a significantly worse fit to the data than a smooth nonparametric model), but other studies find that the nonlinear model fits better. In most applications, linear models for all air pollutants should be an adequate approximation, and always a starting point for more complicated concentration-response relationships.

It would be desirable to have a systematic analytic approach for evaluating the effects of including other pollutants in the PM model. The HEI study illustrates a structured approach. Evaluating all possible copollutant models is the least efficient approach, and is most likely to find spurious confounding by copollutants. The HEI study suggests that in Philadelphia from 1973 to 1988, separating TSP effects from SO_2 effects might require further analysis; that NOx is a weaker confounder, since NOx effects are negative when TSP and/or SO_2 are included and have significant positive effects; that ozone effects are somewhat confounded with the $TSP/SO_2/NO_x$ cluster in summer, but not otherwise; that lagged CO effects are largely independent of other pollutants. EPA has noted that it is difficult to separate the effects of PM from one or another copollutant in certain particular cities (e.g., from SO_2 in Philadelphia). There are, however, studies in other locations where PM effects have been found in the substantial absence of SO_2, such as Los Angeles, Santa Clara, and Utah Valley, which adds credibility to finding a separable PM effect in Eastern cities such as Philadelphia.

Of all of the issues identified by commenters, this copollutant assessment problem is the least fully resolved by epidemiological studies in individual cities with multiple pollutants. At the very least, PM appears to be at the center of a cluster of pollutants that may include gaseous precursors of PM, such as SO_2 and NO_2. For this reason, it is important to examine the evidence across a number of locations. In places where these other pollutants are found at low concentrations, there appear to be distinct effects associated primarily with PM and not attributable to copollutants (mostly absent) or to weather.

It is also important to examine other factors that may provide insights beyond those provided by epidemiological results in specific cities such as Philadelphia. The Staff Paper approaches the issue of separating PM and SO_2 in an integrated assessment that draws on information from atmospheric chemistry, human exposure studies, and respiratory tract penetration results to provide insight as to which of these two pollutants is more likely to be responsible for mortality in the elderly and individuals with cardiopulmonary disease (U.S. EPA 1996b; pp. V-46 to -50). That assessment notes that the inhalable (PM_{10}), including the fine ($PM_{2.5}$), components of TSP are more likely than SO_2 to penetrate and remain indoors where the sensitive population resides most of the time. In addition, these PM components, especially $PM_{2.5}$, penetrate far more effectively to the airways and gas exchange regions of the lung than does SO_2. However, one mechanism by which SO_2 can be transported deeper into the lung is absorption or dissolution onto the surfaces of atmospheric particles. In this case, the complex results reported by HEI in regard to effects associated with SO_2 exposure might be partially reflecting varying atmospheric interactions of the two pollutants, rather than a direct SO_2 effect. Furthermore, in Philadelphia, it

is possible that SO_2 is a surrogate for fine particulate acid sulfates. For these reasons, even though statistical analyses of the Philadelphia data set cannot fully distinguish between these two highly correlated pollutants, EPA believes that the weight of the available evidence from an integrated assessment more strongly supports the notion that PM is playing an important direct role in the observed mortality effects associations in Philadelphia. Moreover, as noted above, in some other locations with significant PM-mortality associations, ambient SO_2 levels are too low to confound PM.

d. Other variables (e.g., weather, day of week)

A number of adequate methods may exist for adjusting health effects models for weather. These include: nonparametric smoothing of temperature and humidity; use of dummy variable categories for ranges of these variables; and use of parametric nonlinear functions such as "U-shaped" temperature response functions (piecewise linear, for example, with increases in response below a cold threshold and above a warm threshold). Interactions may also be worth examining, such as pollution effects on hot and humid days.

A number of inadequate methods have also been used. One of the less defensible proposals is that of a "V"-shaped temperature response function, assuming a unique ideal temperature. Common experience suggests that, for population studies, there is likely to be a range of temperatures which most people find reasonably comfortable; this tends to be confirmed by nonparametric temperature models. It is highly unlikely that the same function would describe different SMSA's populations. Further evaluation may be useful to find an optimal ambient temperature index, perhaps daytime minimum in winter and daytime maximum in summer, rather than mean temperature. Also, several humidity indices have been used by different investigators, including: dewpoint temperature; specific humidity; and relative humidity; determining which (if any) may be more optimal than the others remains to be determined.

Some meteorological variables may be useful in bioclimatology studies, but should likely almost always be avoided in air pollution studies. These include wind speed and barometric pressure, which are likely to be confounders of air pollution, given their likely being good indicators of atmospheric inversion conditions that occur in some seasons and in some locations. Also, wind direction may be a surrogate for conditions when an air pollution monitor is favorably sited to measure emissions from a specific source versus indexing a more general ambient exposure of a study population. The relation or lack of relation of air pollution to certain synoptic weather categories may also involve these variables. In most studies, the independent effects and potential confounding effects of weather appear to be adequately captured by temperature and humidity variables.

EPA notes that the recent HEI re-analyses of Philadelphia have found that a Poisson time series model for mortality produces generally similar estimates of air pollution effects using several different methods of adjustment for time trends and weather, but not for all methods for adjustment. The more flexible methods (data-driven rather than theory-driven) seem to do better

at adjusting for potential confounders. This may not be found in all such studies, but EPA expects that the findings of significant effects of air pollution, and PM specifically where copollutants do not confound the relationship, with PM lag times of 1 and 2 days generally predictive, will be fairly robust for weather and time trends across studies.

Day-of-the-week PM effects have been evaluated in a number of studies and are plausible, but the findings are mixed. In any case, weekend and day-of-the-week effects may be modifying factors of more or less importance, but almost certainly are surrogates for causative factors such as changes in motor vehicle or power plant emissions, and time in indoor or outdoor micro-environments, that reflect weekend-to-work-week transitions.

e. Lags in the pollution and weather variables

Lags have been discussed extensively in the PM CD. There may be real differences in the health effects for different PM lags. Some health responses may have different time scales than other health responses. Because of baseline differences in disease incidence and in population susceptibility in different SMSAs, findings of different time lag structures in different studies may be real differences. Even so, almost all studies that include PM on the two days preceding the day on which the response occurs (lags 1 and 2) find a maximal PM effect. Some studies find that the current day PM (lag 0) is also predictive when included in a moving average concentration, others find lag 0 non-predictive but including lag 3 as well as lags 1 and 2 predictive (for example, the 1993 Schwartz study of Birmingham). When site-specific differences may well exist, the model specification search strategies that evaluate different PM lag structures, starting with lags 1, 1+2 or 1+2+3 as most plausible *a priori*, are justifiable. It is no more sensible to require an absolutely identical PM lag structure in all studies than it is to require an identical weather model in all locations, ignoring local differences in acclimatization to weather. Other pollutants may require other lag times. Ozone effects are likely to be observed on the same day or the following (lags 0 and 1), for example.

Lags for weather effects may be different in different locations, and are very likely different in different seasons at a given location. Nonparametric smoothers may be the best way to capture such differences in a simple model. When pollution episodes are clearly identified with synoptic climatologic categories, these may be sometimes be characterized well by use of the SCC without more complex weather models. The HEI study suggests that when a model for weather and time trends reduces the Poisson overdispersion index to a low value (slightly greater than 1), then little more can be learned by over fitting the weather models.

f. Population or age group affected, cause of death, etc.

Reasonable prior hypotheses are readily evident from the London episodes. During those episodes, mortality occurred primarily in the elderly, from causes that are primarily diagnosed as respiratory or cardiovascular deaths. It is reasonable to hypothesize that similar groups (the elderly, those with preexisting respiratory or cardiovascular conditions, etc.) are likely at risk to lower (non-episode) PM exposures. Some cardiovascular deaths occur in individuals who are apparently healthy enough to leave home even during a very high pollution episode, and so will occur outside of the home or hospital/clinic setting, and are not "death bed" effects. Respiratory diagnoses will occur much more often on high-pollution days than on low-pollution days (i.e. death from respiratory causes has the highest RR), but the association may not be as easily detected as the association with increases in cardiovascular death diagnoses because respiratory deaths occur much less frequently (less than 10 percent of elderly deaths) than do cardiovascular deaths (about 50 percent of elderly deaths).

Very few recent studies have produced findings that disagree with these hypotheses. In some cases, CVD deaths seem to have a higher RR than do respiratory deaths, possibly reflecting little power to identify increases in respiratory RR when the total number of deaths is low (Utah Valley, Steubenville).

In summary, EPA must reject some commenters' contention that alternative analyses can obtain "almost any result". Strong arguments can be made for the methods and analytical strategies defined in this Response as more probably legitimate approaches than other approaches. The burden of proof on what constitutes a "legitimate" model or analytical strategy rests on those who propose approaches that deviate substantially from the approaches described above. This would have to be evaluated on a case-by-case basis, at least to the level of detail in the PM CD.

Comment: Some commenters have argued that, given the flexibility in modeling strategies addressed in the previous comment and the pressure to publish positive results, there exists considerable uncertainty about the uniqueness of potential conclusions, and that legitimate alternative methods used to date have led to alternative conclusions.

RESPONSE: EPA's concerns about the effects of using alternative methods have led to extensive evaluations of methodology and analytical strategy, starting with discussions at the California PM Workshop at UC-Irvine in Jan. 1994 and the Mortality Workshop sponsored by EPA in RTP, NC, in November, 1994. While commenters note the possible pressure to publish "positive" findings, there was actually an enormous increase in the rate of publication of both "positive" and "negative" findings in the last few years. No synthesis of the diverse findings was feasible without a detailed evaluation of each individual study and the methods used in arriving at the study's findings.

The discussion of methodology in the earliest drafts of the PM CD was much more extensive than in the final draft, reflecting the assessment of CASAC that investigators were beginning to

focus on a much more limited range of appropriate methods. Some investigators have continued to use methodology quite outside this range, and CASAC as well as EPA staff have pointed out deficiencies in some of the alternative methods. EPA's conclusion is that the range of "legitimate" methods is quite a bit narrower than used in some of the studies cited by commenters, which have been given less weight because of their inadequate methodology. See comments on such methodologic inadequacies presented elsewhere (Appendix C) under responses concerning more recent individual studies.

Comment: Dr. Ron Wyzga of EPRI claimed that he and colleagues have found significant air pollution-health effects associations even when doing time series analysis with data sets that should have absolutely nothing to do with each other. More superficially, they indicated finding such results if they chose the optimal lag in daily Philadelphia ozone levels (lagged 1 through 4) and daily mortality in Santa Clara County for 1985. They then posed the question: could it be that there is an underlying artifact in large seasonally-varying data sets that can often lead to positive associations when such associations are not warranted?

RESPONSE: Beyond these assertions made in public testimony by Wyzga and in his comments for EPRI, no documentation describing the analyses alluded to has been provided to EPA for consideration, either in the form of a published or accepted paper or even as a manuscript describing the methods, data, and results. It is obvious that, since these analyses have not been presented to EPA for review, it is not possible to make a detailed technical evaluation but some relevant questions can be posed here. Given the results are reported as having been found in specific years, 1985, 1988, and 1992, were these years selected at random, selected for significant findings after all years had been analyzed, or in some other way? How were the other variables selected? How stable were the results for one year versus another?

Comment: A number of commenters stated that we don't know if the differences in mortality across the 6 cities included in the Dockery, et al. (1993) study are due to an uncontrolled regional variable (sedentary life style), to differences in air quality, to some combination of these two, or perhaps to neither variable. In cross-sectional designs other pollutants can be confounders as well. Lipfert (1995) has shown, using statewide data from the National Center for Health Statistics that this variable fits the regional differences in mortality across the 6-cities as well as any air quality variable.

RESPONSE: EPA was aware of the Lipfert (1995) study at the time of PM CD preparation and CASAC review. Individual-level data on "sedentary life style" may be useful under some circumstances, but this indirect measure of potential cardiovascular risk contributing factor can hardly be argued to be better than the use of more direct cardiovascular risk factor data (e.g., body mass index). While 'sedentary lifestyle' is a candidate for inclusion as an a covariate in individual risk factor studies, such as prospective cohort studies, it is not necessarily appropriate for inclusion as a covariate in cross-sectional "ecologic" studies involving specific cities, since within state regional differences in this variable are also possible. The availability of only the

highly aggregated statewide level from NCHS was regarded as probably much noisier than regional air pollution data for PM_{10}, $PM_{2.5}$, and sulfates.

Differences in "sedentary lifestyle" may also exist across different regions or cities within a given state, such that without verifying that a given city matches the state average, then the state-wide value assigned for a given city could be spuriously high or low and thereby alter the purported "good fit" to the Harvard Six-Cities study data reported by Lipfert (1995).

Comment: Some commenters also asserted that, since many pollutants are highly inter-correlated in the ambient environment, it is difficult (and often impossible) to indicate which pollutant (including particulate matter) is associated with a health endpoint.

RESPONSE: One of the most important concerns is that, even when there is a definite health effect of air pollution exposure, this effect may not always be assignable to PM or to some other single pollutant. Pollutant concentration correlations, as noted in the PM CD, have many causes: weather conditions, common sources (e.g. motor vehicles), and some pollutants are precursors to PM components (e.g., sulfates, nitrates). Multiple causal factors are frequently encountered in epidemiology and in other fields, and there are a variety of analytical approaches to untangling them. Principal components analysis/factor analysis may allow a considerable reduction in the dimensionality of the problem, distinguishing those common factors that can be substantially separated from other factors. For example, EPA reanalyses of the Samet (1996a,b) report on multiple pollutants in Philadelphia suggested that there was a general factor involving TSP, SO_2, NO_2, and CO. There was also considerable separation of O_3 from this component except in summer, and another distinct component composed primarily of TSP and SO_2. A second approach discussed at the PM Mortality Workshop in Nov. 1994 was causal (structural) modeling, disentangling the components of the co-pollutant structure based on source profiles.

The only assessment of potential confounding of a possible PM effect by co-pollutants has been based on the sensitivity of the estimated PM regression coefficient to inclusion of other pollutants. Not surprisingly, most studies found substantial sensitivity, typically a reduction of the PM coefficient although this has not always been the case. It is also not surprising, in view of the air pollutant correlations, that adding one or more correlated pollutants to a regression model can greatly increase the estimated uncertainty of the PM coefficient (standard error) and thereby decrease its statistical significance *even when the estimated size of the PM effect is virtually unchanged*.

Some investigators such as HEI have explored this issue in detail using two-dimensional nonparametric smoothing surfaces (Samet et al., 1995). The problem may be more complex than it looks, since this includes using air pollution metrics with the appropriate lag or moving average structure, and pollution effects may change by season. For PM, the London studies show that there are significant effects from BS (PM) exposure on the two previous days and a smaller effect for current-day exposure, whereas ozone effects seem to described best by same-day or previous-day concentrations.

Goodness-of-fit assessments against nonparametric smooth functions of air pollution are feasible, but rarely done. The few analysts who have pursued these model specification tests tend to find that a linear or piecewise linear model for PM_{10} is often adequate (i.e., not providing a significantly worse fit to the data than a smooth nonparametric model), but other studies find that the nonlinear model fits better. In most applications, linear models for all air pollutants should be an adequate approximation, and always a starting point for more complicated concentration-response relationships.

It would be desirable to have a systematic analytic approach for evaluating the effects of including other pollutants in the PM model. The HEI study illustrates a structured approach. Evaluating all possible copollutant models is the least efficient approach, and is most likely to find spurious confounding by copollutants. The HEI study suggests that in Philadelphia from 1973 to 1988, separating TSP effects from SO_2 effects might require further analysis; that NOx is a weaker confounder, since NOx effects are negative when TSP and/or SO_2 are included and have significant positive effects; that ozone effects are somewhat confounded with the $TSP/SO_2/NO_x$ cluster in summer, but not otherwise; that lagged CO effects are largely independent of other pollutants.

A substantial effort has been made to identify specific pollutants or groups of pollutants that are associated with health endpoints in various studies. While complete separation of effects from different pollutants may not be possible in all studies, much progress has been made in this direction. As noted above, although it is difficult to separate PM effects from those of one or another pollutants in the same city in some cases, PM effects are separable from the same pollutant in other cities or remain positive and significant regardless of whether a given specific copollutant is present in high or low concentrations in various other locations. The basis is less clear for determining whether health effects are associated with PM concentrations because of the number of particles, the chemical constituents of PM, or the interaction of PM with other air pollutants. See also earlier responses relevant to the effects of copollutants.

Comment: Some commenters take issue with EPA's arguments that studies undertaken at various locations in various seasons where some pollutants occur only at very low levels, adequately address the multiple pollutant issue. For example, Wyzga notes that for Santa Clara County in the winter, where SO_2 is low and ozone is not likely the culprit, data were also available for carbon monoxide (CO). When CO levels were substituted for particulate measures they were significantly associated with daily mortality in Santa Clara, based on an unpublished analysis by Lipfert.

RESPONSE: The study by Lipfert has not been made available for review by EPA, so no full specific technical responses are possible. It is entirely possible that statistical correlations among pollutants may exist at levels that are so low as to make adverse health effects from one of the pollutants highly unlikely. The 1990 Santa Clara PM index is in Coefficient of Haze (COH) units that do not correspond precisely to PM_{10}, although Fairley (1990) cites approximate ratios of BS/COH = 0.55 and TSP/COH = 1. If CO levels are much below the NAAQS, then CO

levels may be well correlated with COH levels because both pollutants are combustion by-products, but the CO levels may be so low that it is unlikely that CO can be causing excess deaths. Since EPRI did not report levels for CO used in these analyses, it is not possible to adequately review these reported Lipfert findings.

Even assuming that the unpublished Lipfert analyses were correct, PM contributions to mortality were separable from those of CO in Philadelphia (Samet, 1996 a,b; EPA 1996a) and there is no more basis for believing CO to be confounding the PM results in Santa Clara. Moreover, EPRI ignores comparisons in other locations where PM is significant in the presence of high or low levels of other specific copollutants. Also, EPA notes that another commenter (the Bay Area Air Quality control District) submitted an analysis extending the original Santa Clara mortality study. This reanalysis identifies a close relationship between $PM_{2.5}$ and COH in this location.

Comment: Various commenters also noted that presence of exposure error can bias the measure of association and the shape of any exposure-response curve and that the consequences are even more serious in a multivariate context.

RESPONSE: As discussed more extensively in appendix D, exposure measurement error issues will have little effect on the use of epidemiology models for estimation of health effects or other risk assessment applications. A very comprehensive summary of modern approaches to statistical inference in regression models when measurement error is present in the predictors has been published by Carroll et al. (1995). They conclude:

> "If a predictor X is measured with error, and one wants to predict a response *based on the error-prone version W of X*, then ... it rarely makes sense to worry about measurement error. ... The one situation requiring that we model the [distribution of] measurement error occurs hen we develop a prediction model using data from one population, but we wish to predict in another population. A naive prediction model that ignores measurement error may not be transportable." (Carroll, et al. 1995)

The complete characterization of exposure measurement error includes the exposure measurement error structure of important covariates or modifying factors, such as weather. Since this is likely to vary significantly from place to place, predictive models for one SMSA that include weather-related factors may not be transportable to some other SMSA. However, to the extent that coefficients for PM_{10} and $PM_{2.5}$ appear to be more similar for different SMSAs than do coefficients for weather variables, the estimated effects for PM factors may be reasonably transportable. See more detailed measurement error discussion in Appendix D.

Comment: A number of commenters relied on Lipfert and Wyzga (1995) in commenting on the relative importance of PM or PM constituents effects versus those of other pollutants or PM constituents. Lipfert and Wyzga (1995) reviewed the extensive literature and demonstrated that there was little difference in the elasticity (measure of association used; see below) among the various particulate measures. More specifically, when they compared elasticity for mortality and

PM_{10} with the elasticity for mortality and $PM_{2.5}$, there was little difference with the elasticity for PM_{10} being larger than that for $PM_{2.5}$. In public comments, Wyzga gives the following definition of "mean effect" or elasticity:

> "The 'mean effect' or elasticity at the mean is a quantitative continuous number which describes the relationship between an endpoint and a risk factor. This measure is obtained by multiplying the regression coefficient by the mean concentration of the pollutant. The mean effect takes into account all of the excursions of a pollutant weighted by the frequency of occurrence and estimates how the endpoint would change if the risk factor were completely eliminated." [EPRI, IV-D-2672]

RESPONSE: The above quote by Wyzga attempts to correct a fundamental problem in the paper by Lipfert and Wyzga (1995), the mathematical misspecification of the concentration-response function. For various reasons specified below, EPA finds that this elasticity approach is inappropriate for making the kinds of comparisons that Wyzga and Lipfert do between relative contributions of PM, different PM subcomponents, or other copollutants. The PM CD evaluated a number of possible specifications of the mathematical functions that might be plausible representations of the relationship between PM and mortality.

Lipfert and Wyzga (1995) also discuss the results from 31 epidemiology studies which have shown associations between increased mortality and air pollution. In an attempt to compare the studies in a manner which is independent of measurement units and averaging times, elasticity was used as the measure of effects. They compared the elasticities of two basic study types: time-series studies and cross-sectional studies. For time-series regression, they assert that the results vary according to the method of seasonal adjustment, covariates included and lag structure assumed. For cross-sectional studies, they claim that the results are highly dependent on methods used to control for socio-economic and personal lifestyle factors and on data quality. The authors contend that the studies are unable to differentiate the response of collinear pollutants and weather. Based on the lack of statistical difference among responses to the various particle measures (TSP, PM_{10}, $PM_{2.5}$, COH, or $SO_4^=$) in the various studies, Lipfert and Wyzga claim that there is no single responsible pollutant. They conclude that all of the pollutants may be potential culprits in premature mortality in severely compromised patients.

It is important to point out that the main premise on which the Lipfert and Wyzga (1995) review is based is inherently flawed. The authors have based their entire synthesis of the 31 studies on an exposure metric which is inappropriate for the purposes of this review. Elasticity, as discussed later, does not allow a fair comparison of risk estimates between studies. With this in mind, it should made clear that trends in magnitude and significance of different particulate matter risk estimates have in fact been demonstrated in various studies. One such study (omitted by Lipfert and Wyzga) is that of Ozkaynak and Thurston (1987), in which a cross-sectional analysis of 100 U.S. standardized metropolitan statistical areas demonstrated that particle size and composition were important factors in modeling particle pollution health effects. Overall, Ozkaynak and Thurston found that the PM coefficients (and elasticities) had the following

ranking: TSP < Inhalable Particles < Fine Particles < $SO_4^=$. Similar conclusions were reached by Thurston et al. (1994). Thus, a more objective inclusion of all of the relevant studies might lead to different conclusions than those presented in this paper.

The elasticity of a regression coefficient is defined as the percentage change in a dependent variable (such as mortality) associated with a 1% change in an independent variable (such as PM air pollution). As discussed in the Lipfert and Wyzga (1995) paper, the elasticity has the characteristic that it can allow dimensionless intercomparisons of adverse effects across various pollutants. However, in making this normalization to derive elasticity estimates, the absolute size of the effect (e.g., in numbers of deaths per concentration of pollutant) is lost, which is an important part of the intercomparison. Although convenient for interpollutant comparisons at a single site, the basic problem with using the elasticity is that the health effects of air pollution have been found (e.g., in exposure chamber studies) to increase in proportion to the absolute concentration of the pollutant, not in proportion to the percent increase in pollution. For example, the change in effects in going from 50 to 100 $\mu g/m^3$ would be expected to be the same as for an increase from 100 to 150 $\mu g/m^3$ but the elasticity would indicate that the former is a 100% increase, while the latter would be only a 50% increase. Thus, if the PM effect coefficients (in deaths/$\mu g/m^3$) were identical in cities A and B, but the average pollution levels in City A were double those in City B, the elasticity in city A would be half that in city B, which would be misleading. In other words, in order to inter-compare properly these elasticities across cities with differing pollution levels, they must be converted back to absolutes. Similarly, differing cities have differing underlying rates of adverse health effects (e.g., of death per 100,000, which depends on varying factors such as the population's age distribution), and a given percentage change in health outcome would also have a different meaning in each city. Thus, the concept of elasticities is really not the most appropriate for the study to study (and usually city to city) intercomparison purposes Lipfert and Wyzga propose in their paper.

To emphasize, as noted above, the "elasticity" representing the percentage increase in deaths per 1% increase in PM. This implies that the effect of increasing exposure concentration from 1 $\mu g/m^3$ to 2 $\mu g/m^3$ is exactly the same as that of increasing exposure concentration from 10 $\mu g/m^3$ to 20 $\mu g/m^3$, or from 100 $\mu g/m^3$ to 200 $\mu g/m^3$, or from 1000 $\mu g/m^3$ to 2000 $\mu g/m^3$. EPA believes this is highly implausible. While some studies such as the London daily mortality time series analyses (1986 PM/SO_X CD) showed a flatter slope at BS concentrations above 1000 $\mu g/m^3$, the relationship appeared sensibly linear below 250 to 500 $\mu g/m^3$. Nonparametric RR smoothing functions fitted by various investigators are often visually and statistically indistinguishable from linear functions in the range 20 or 30 to 150 $\mu g/m^3$, with too few data to identify the concentration-response function outside this range. The Linear and log-linear models are almost indistinguishable at current PM levels, and show little similarity to the flat downward-curving shape of the log-log models.

In summary, the representation of concentration-response relationships in the PM CD and in other papers cited by EPA, including the paper by Schwartz et al. (1996) on fine and coarse particle effects, provides a much more reliable basis for inference and extrapolation. The use of

"elasticity" or "mean effect" distorts the real relationship and ignores the shape of the exposure-response curve. Comparisons using dimensionless indices are of little use in making distinctions between relative contributions of PM, its various subcomponents, or other copollutants, where "how big" and "how much" are relevant questions.

III. KEY PM EPIDEMIOLOGICAL STUDIES IDENTIFIED IN CHAPTER 12 OF THE CRITERIA DOCUMENT

A number of commenters on the EPA's proposed PM standards raised questions regarding what key studies EPA used in its evaluation of the scientific literature. Further questions and assertions have been raised with respect to the numbers of studies finding associations between PM and various health endpoints. In essence, EPA has relied on the scientific information contained in the PM CD and Staff Paper, and with respect to the epidemiological studies, those studies discussed and summarized in Chapter 12. For clarification, the key studies in Chapter 12 forming the major basis for EPA's findings on the consistency and coherence of the epidemiological data have been summarized in the attached table. Of course, the EPA's decision does not rest solely on findings from epidemiological studies, but also on results from studies of other aspects of the scientific information in the CD, such as exposure, physico-chemical characteristics, dosimetry, and toxicology. A number of epidemiological studies that are not included in this table, including cross-sectional studies, were also reviewed in the CD and provided additional information for the comprehensive weight-of-evidence evaluation of the scientific literature.

The key studies were derived from the following summary tables in Chapter 12 in the CD: Tables 12-2, 12-8, 12-9, 12-10, 12-11, 12-12, 12-13, 12-16, 12-21, and 12-22. These summarize results and characteristics of epidemiological examinations of PM exposure and short-term mortality, long-term mortality, hospitalization and emergency room admissions, short-term or long-term changes in lung function or prevalence/incidence of symptoms or medication use. The studies are presented in the attached table as they were grouped for presentation in the CD, and each group is identified by the appropriate table from the CD. A total of 87 different studies are included. In the CD, single studies with multiple effects sometimes were listed in more than one table, which were organized by effects categories.

Based on the CD summaries and the studies themselves, staff categorized the PM-effects associations from each study as "significant effects", "mixed", or "no significant effects". A study is categorized as having significant effects if a statistically significant association is reported between one or more health effects and one or more indicators of PM. This group of studies includes a number in which all PM-health associations reported are statistically significant, such as Pope et al. (1992), where significant positive associations are found between PM_{10} and mortality from all causes, respiratory and cardiovascular diseases. The "mixed" category includes studies where results with respect to PM effects are less clear, including, for example, multi-pollutant studies in which the authors noted difficulties in separating the effects

of PM from other pollutants, even if PM is significant (e.g. Samet et al., 1996a,b; Moologavkar et al., 1995b). The "no significant associations" category includes those studies testing PM associations, alone or in combination with othernpollutants, that reported no significant, nearly significant, or mixed results between PM concentration and the health endpoints measured. It should be noted that none of the studies in the table report finding consistent, statistically significant negative associations between PM and adverse effects.

As shown in this summary Table, the vast majority of the studies had statistically significant associations between PM indicators and health effects; 68 (78%) reported statistically significant associations, 11 (13%) studies had "mixed" findings and 8 (9%) studies report no statistically significant results.

TABLE B-1. KEY EPIDEMIOLOGICAL STUDIES REGARDING PM EXPOSURE AND HEALTH EFFECTS

PM Measure	Location	Reference	Significant effects	Mixed effects	No significant effects
SHORT-TERM PM EXPOSURE (1 to 5 Days) -- MORTALITY (from CD, Table 12-2)					
KM	Los Angeles	Shumway et al. (1988)	1	0	0
KM	Los Angeles	Kinney & Ozkaynak (1991)	0	1	0
COH	Santa Clara, CA	Fairley (1990)	1	0	0
BS	London	Thurston et al. (1989)	1	0	0
BS	London	Ito et al. (1993)	1	0	0
Suspended Particles	Erfirt, Germany	Spix et al. (1993)	1	0	0
BS	Athens, Greece	Katsouyanni et al. (1990a)	1	0	0
BS	Athens, Greece	Katsouyanni et al. (1990b)	0	1	0
BS	Athens, Greece	Katsouyanni et al. (1993)	0	0	1
BS	Athens, Greece	Touloumi et al. (1994)	1	0	0
TSP	Detroit	Schwartz (1991a)	1	0	0
TSP	Philadelphia	Schwartz & Dockery (1992a)	1	0	0
TSP	Philadelphia	Moolgavkar et al. (1995b)	0	1	0
TSP	Philadelphia	Wyzga & Lipfert (1995b)	0	1	0
TSP	Philadelphia	Li & Roth (1995)	0	1	0
TSP	Philadelphia	Samet et al. (1995)	1	0	0
TSP	Philadelphia	Samet et al. (1996b)	0	1	0
TSP	Philadelphia	Samet et al. (1996a)	0	1	0
TSP	Philadelphia	Cifuentes & Lave (1996)	1	0	0
TSP	Steubenville, OH	Schwartz & Dockery (1992b)	1	0	0
TSP	Steubenville, OH	Moolgavkar et al. (1995a)	0	1	0
TSP	Cincinnati, OH	Schwartz (1994a)	1	0	0
TSP	Lyons, Marseilles, France	Derriennic et al. (1989)	0	0	1
TSP	Beijing, China	Xu et al. (1994)	1	0	0

PM_{10}	Utah County	Pope et al. (1992)	1	0	0
PM_{10}	Utah County	Pope & Kalkstein (1996)	1	0	0
$PM_{2.5}$, H^+, PM_{10}	St. Louis, MO; Kingston, TN	Dockery et al. (1992)	1	0	0
PM_{10}	Birmingham, AL	Schwartz (1993a)	1	0	0
PM_{10} (est,) TSP,COH	Toronto	Ozkaynak et al. (1994)	1	0	0
PM_{10}	Los Angeles	Kinney et al. (1995)	1	0	0
PM_{10}	Los Angeles, Chicago	Ito et al. (1995)	1	0	0
PM_{10}	Santiago, Chile	Ostro et al. (1996)	1	0	0
PM_{10}	Sao Paulo, Brazil	Saldiva et al. (1994)	0	0	1
PM_{10}	Sao Paulo, Brazil	Saldiva et al. (1995)	1	0	0
PM_{10}	Cook Cnty, IL; Salt Lake Cnty	Styer et al. (1995)	0	1	0
PM_{10}	Utah County	Lyon et al. (1995)	0	0	1
PM_{10}	Cook County, IL	Ito & Thurston (1996)	1	0	0
PM_{10}, $PM_{2.5}$, CP, H^+, sulfate	6 Cities	Schwartz et al (1996)	1	0	0
		Mortality Studies Subtotal	**24**	**10**	**4**
SHORT-TERM PM EXPOSURE (1 to 5 Days) -- HOSPITAL ADMISSIONS/VISITS FOR RESPIRATORY CAUSES (CD, Table 12-8)					
Sulfate	Ontario, Canada	Burnett et al. (1994)	1	0	0
Sulfate, TSP, PM_{10}, $PM_{2.5}$, CP, H^+	Ontario, Canada	Thurston et al. (1994a,b)	1	0	0
Sulfate, H^+	Buffalo, Albany, NY	Thurston et al. (1992)	1	0	0
PM_{10}	New Haven, CT; Tacoma, WA	Schwartz (1995a)	1	0	0
PM_{10}	Cleveland	Schwartz et al. (1996)	1	0	0
PM_{10}	Spokane, WA	Schwartz (1996)	1	0	0
PM_{10}	Seattle	Schwartz et al. (1993)	1	0	0
PM_{10}	Tri-cities, WA	Hefflin et al. (1994)	1	0	0
PM_{10}	Anchorage, AK	Gordian et al. (1996)	1	0	0

SHORT-TERM PM EXPOSURE (1 to 5 Days) -- HOSPITAL ADMISSIONS/ER VISITS FOR COPD (CD, Table 12-9)					
BS	Barcelona, Spain	Sunyer et al. (1993)	1	0	0
PM_{10}	Birmingham, AL	Schwartz (1994e)	1	0	0
PM_{10}	Minneapolis	Schwartz (1994f)	1	0	0
PM_{10}	Spokane, WA	Schwartz (1996)+	1	0	0
PM_{10}	Detroit	Schwartz (1994d)	1	0	0
SHORT-TERM PM EXPOSURE (1 to 5 Days) -- HOSPITAL ADMISSIONS/ER VISITS FOR PNEUMONIA (CD, Table 12-10)					
PM_{10}	Minneapolis	Schwartz (1994f)+	1	0	0
PM_{10}	Birmingham	Schwartz (1994e)+	1	0	0
PM_{10}	Detriot	Schwartz (1994d)+	1	0	0
PM_{10}	Spokane, WA	Schwartz (1996)+	1	0	0
TSP	Philadelphia	Schwartz (1994g)	1	0	0
SHORT-TERM PM EXPOSURE (1 to 5 Days) -- HOSPITAL ADMISSIONS/ER VISITS FOR HEART DISEASE (CD, Table 12-11)					
PM_{10}	Detroit	Schwartz & Morris (1995)	1	0	0
Sulfate	Ontario, Canada	Burnett et al. (1995)	1	0	0
		Hospital Admissions Totals	21	0	0
		Hospital Admissions Subtotals (minus double counts)	**16**	**0**	**0**

SHORT-TERM PM EXPOSURE (1 to 5 Days) -- PREVALENCE/INCIDENCE OF RESPIRATORY SYMPTOMS (CD, Table 12-12)					
PM_{10}, $PM_{2.5}$, $PM_{2.5}$/S, H^+	Six U.S. Cities	Schwartz et al. (1994)	1	0	0
PM_{10}	Utah Valley	Pope et al. (1991)	1	0	0
PM_{10}	Utah Valley	Pope & Dockery (1992)	1	0	0
PM_{10}, BS, H^+	Wagerningen, Netherlands	Hoek & Brunekreef (1993)	1	0	0
TSP	Five German Communities	Schwartz et al. (1991a)	1	0	0
TSP	Four Switzerland Cities	Braun-Fahrlander et al. (1992)	1	0	0
PM_{10}, BS, H^+	Wageningen, Netherlands	Roemer et al. (1993)	1	0	0
PM_{10}	1 Netherlands Community	Dusseldorp et al. (1994)	1	0	0
$PM_{2.5}$, H^+, nitrate, sulfate, nitric acid	Denver	Ostro et al. (1991)	1	0	0
Sulfate, COH	Southern Californina	Ostro et al. (1993)	1	0	0
PM_{10}	Los Angeles	Ostro et al. (1995)	1	0	0
Sulfate, Nitrate, PM_{10}	Netherlands (300 subjects)	Hoek & Brunekreef (1995)	0	0	1
$PM_{2.5}$, PM_{10}, sulfate, PSA	Uniontown, PA (83 subjects)	Neas et al. (1995)	1	0	0
PM_{10}, sulfate, nitrate, HONO	Four Netherlands Cities	Hoek & Brunekreef (1994)	1	0	0
		Respiratory Symptoms Subtotals	**13**	**0**	**1**

SHORT-TERM PM EXPOSURE (1 to 5 Days) -- REDUCED LUNG FUNCTION (from CD, Table 12-13)					
TSP	Steubenville, OH	Dockery et al. (1982)	1	0	0
TSP, RSP (PM_{10})	Ijmond, Netherlands	Dassen et al. (1986)	1	0	0
$PM_{2.5}$, PM_{10}	Tuscon, AZ (6 subjects)	Quackenboss et al. (1991)	0	0	1
PM_{10}	Utah Valley	Pope et al. (1991)+	1	0	0
PM_{10}	Utah Valley	Pope & Dockery (1992)+	1	0	0
$PM_{2.5}$	Seattle	Koenig et al. (1993)	1	0	0
BS, PM_{10}, BS, H^+	Wageningen, Netherlands	Hoek & Brunekreef (1993)+	1	0	0
BS, PM_{10}, BS, H^+	Netherlands	Roemer et al. (1993)+	1	0	0
PM_{10}	Utah Valley	Pope & Kanner (1993)	1	0	0
$PM_{2.5}$, PM_{10}, sulfate, PSA	Uniontown, PA (83 subjects)	Neas et al. (1995) +	1	0	0
PM_{10}, sulfate, H^+, ammonium	Austria (133 subjects)	Studnicka et al. (1995)	0	1	0
PM_{10}, sulfate, nitrate, HONO	Four Netherlands Cities	Hoek & Brunekreef (1994) +	1	0	0
PM_{10}	Wijkaan Zee, Netherlands	Dusseldorp et al. (1994)+	1	0	0
		Lung Function Totals	11	1	1
		Lung Function Subtotals (minus double counts)	**4**	**1**	**1**
LONG-TERM STUDIES: PROSPECTIVE COHORT MORTALITY (CD Table 12-16)					
TSP	California Communities	Abbey et al. (1991a)	0	0	1
PM_{15}, $PM_{2.5}$, Sulfate	Six U.S. Cities	Dockery et al. (1993)	1	0	0
$PM_{2.5}$, Sulfate	47/151 U.S. Cities	Pope et al. (1995b)	1	0	0
LONG-TERM STUDIES: CHRONIC RESPIRATORY DISEASE (CD Table 12-21)					
TSP	Six U.S. Cities	Ware et al. (1986)	1	0	0
PM_{15}	Six U.S. Cities	Dockery et al. (1989)	1	0	0
PM_{10},$PM_{2.1}$, sulfate,acidity	24 U.S. Cities	Dockery et al. (1996)	1	0	0
TSP, PM_{10}	California Communities	Abbey et al. (1995a,b,c)	3	0	0

LONG-TERM STUDIES: CHRONIC PULMONARY FUNCTION CHANGE (CD Table 12-22)					
TSP, Sulfate	Six U.S. Cities	Ware et al. (1986)+	0	0	1
PM_{15}, $PM_{2.5}$, Sulfate	Six U.S. Cities	Dockery et al. (1989)+	0	0	1
$PM_{2.5}$, Sulfate	Six U.S. Cities	Neas et al. (1994)	0	0	1
$PM_{2.1}$, PM_{10}, Sulfate, PSA	24 U.S. Cities	Raizenne et al. (1996)	1	0	0
$PM_{2.5}$,PM_{10}	Cubatao, Brazil	Spector et al. (1991)	1	0	0
TSP, PM_{10}	8 Switzerland Areas	Ackermann-Liebrich et al. (1996)	1	0	0
		Total Long-Term Studies	11	0	4
		Long-Term Subtotal (minus double counts)	**11**	**0**	**2**
	Short-term Study Summary	**Mortality Study Subtotals**	**24**	**10**	**4**
		Hospital Admissions Subtotals	**16**	**0**	**0**
		Respiratory Symptoms Subtotals	**13**	**0**	**1**
		Lung Function Subtotals	**4**	**1**	**1**
	Long-Term Study Summary	**Total, All Long-Term Subtotals**	**11**	**0**	**2**
	ALL STUDY SUMMARY		**68**	**11**	**8**

NOTE: BS = black smoke
PSA = particle-strong acidity
TSP = total suspended particles

APPENDIX C. STUDIES PUBLISHED AFTER COMPLETION OF THE CRITERIA REVIEW

As discussed in section II.B of the preamble to the final notice, a number of epidemiological or other studies were cited or submitted by commenters opposed to revisions of the standards were published or otherwise made available only after CASAC closure on the the PM Criteria Document and Staff Paper. EPA agrees that it did not rely on these studies, based on its long-standing practice of basing NAAQS decisions on studies and related information included in the pertinent air quality criteria and available for CASAC review. The Agency nevertheless has conducted a provisional examination of these and other recent studies to assess their general consistency with the much larger body of literature evaluated in the Criteria Document and Staff Paper; the examination is presented below in four sections. Section I assesses the studies most commonly submitted or cited by commenters as providing results that are contradictory to the generally consistent body of evidence cited by EPA as supporting the need to revise the primary PM NAAQS. Among those most frequently cited are reanalyses by Davis, et al. (1996), Roth and Li (1997), Moolgavkar, et al. (1997), and Lipfert and Wyzga (1997); the latter paper deals, in part, with some overarching methodological issues and concerns of broad relevance for PM epidemiology analyses. Commenters cited still other new analyses of data from other countries, including analyses by: Morris, et al. (1996), for Manchester, UK; Burnett et AL. (1997) for Toronto, Canada; Roth (1997) for Prague, Czech Republic; and one or more studies that were conducted as part of the European multinational "APHEA" project. Because the APHEA studies encompass a larger body of studies than the more limited number these commenters noted, Section II of this Appendix summarizes the extensive European APHEA work as a coherent, interrelated set of studies to provide a more representative examination. Section III presents the result of a limited search and provisional assessment of several other newly available community epidemiological studies identified by EPA but not cited by commenters, followed by brief summarization of some new toxicological work in Section IV.

EPA notes that the assessment presented here is much less inclusive and rigorous than a criteria review. Although these studies were not considered in the final decision on the PM standards, EPA believes this provisional assessment calls into question commenters' assertion that full consideration of selected new studies in the final decision would materially change the Criteria Document and Staff Paper conclusions on the consistency and coherence of the PM data, or on the need to revise the current standards.

I. NEWLY AVAILABLE EPIDEMIOLOGICAL STUDIES CITED BY COMMENTERS AS NOT SUPPORTING EPA'S DECISION

A. Review of Davis, et Al. (1996) Report on Airborne Particulate Matter and Daily Mortality in Birmingham, Alabama (NISS Tehnical Report #55).

Several commenters alluded to one or more new reanalyses of data from Birmingham, Alabama. One of the most frequently cited analyses was an unpublished study by Davis et al. (1996) described in a National Institute of Statistical Sciences (NISS) report, said by some commenters to have been done under contract to EPA. The Davis et al. analysis tries to ensure that the effect of hot and humid days is considered. The key conclusion stated by Davis, et al. and cited by commenters: "When we use the same variables as included by Schwartz, we obtain similar results to his ... when humidity is included among the meteorological variables (it is excluded in the analysis by Schwartz [1993]), we find that the PM_{10} effect is not statistically significant." At the time that the PM CD was closed, the Davis et al. (1996) study had not yet been written and so was not available for review by EPA in the PM CD. However, the paper is now reviewed below, including the recent additional analyses by the investigators (beyond those in the NISS report) that actually verify originally reported positive findings.

The first thing to note is that several factual errors were made in the Davis et al. (1996) study, including:

(1) the variables used by Davis et al. differed from Schwartz (1993) in that the PM_{10} metric used by Schwartz was the average of PM_{10} concentrations on the three days preceding the day of death, whereas the PM_{10} metric used by Davis et al. was the average of the PM_{10} on the day of death and the two preceding days. Since the current-day PM is often not predictive or even negatively predictive of death, this is clearly not the same model as used by Schwartz (1993). Further analysis by one of the co-authors (R. Smith, personal communication, Feb. 8, 1997) showed that when Schwartz's PM metric was used, the estimated PM_{10} effect was of about the same magnitude, but statistically significant at the 0.05 level, even using the very different characterization of humidity effect used by Davis et al.;

(2) Schwartz did include humidity in his 1993 study, and his finding of a hot-and-humid day effect was reported there;

(3) Davis et al. also used a different humidity indicator, specific humidity (absolute water content of the air). It therefore appears that Davis et al. were, in fact, able to replicate the original PM_{10} results reported by Schwartz, and they were robust against a very different weather model specification. Finally, NISS was not under contract to the EPA office that supported this study when the subject report was completed (Nov., 1996).

Additional discussion related to these and other points is provided below.

Schwartz reported results based on an appropriate model specification test: "The mean of PM_{10} concentrations on the 3 previous days had a higher model [chi-squared] than shorter or longer averaging periods ...". (P 1141, C2, L6-8). The PM index use in the Davis et al. report is different than Schwartz's mean of the three preceding days, and has been found by other investigators (e.g. Roth and Li, 1996) to be much less predictive than the PM index evaluated and used by Schwartz. The current-day effect in such models is often negative, and including the current-day PM concentration in a moving average may substantially diminish the estimated PM effect in Birmingham. In Birmingham, there is reason to believe that excess mortality is associated with approximately a two-day lag or with the PM moving average over the preceding three days, excluding the current day's concentration. This was confirmed by these investigators in their March 12, 1997 submission to EPA: "However, the results in [Davis et al., 1996, under revision; Schwartz, 1993] show a significant effect when the three-day average EXCLUDES the current day, but not when it INCLUDES the current day. ... we have examined the effects of individual days' contributions of PM_{10}, producing a statistically significant NEGATIVE effect for the current day, and significant positive effects for the one-day and three-day lags. ..." These results tend to verify that the Schwartz (1993) findings are robust against different model specifications and are clearly different than the impression to the contrary left by the prematurely circulated initial NISS Report 55 authored by Davis et al.

Davis, et al. also mistakenly assert "it would appear that the only meteorological variables actually used by Schwartz were daily mean temperature and an indicator of whether daily mean temperature was over 28 [degrees] C, both lagged one day." However, Schwartz (1993) states clearly: "the initial analyses of weather dependence considered 24-hour mean temperature and dewpoint temperature, a dummy variable for hot days, a dummy variable for humid days, a dummy variable for cold days, and a dummy variable for hot and humid days. Weather variables on the concurrent day and on the previous 3 days were considered. The definitions of hot, humid, and cold were set at the 95th percentile .. weather terms were kept in the regression model as long as they were at least marginally significant ($t > 1$)." The statistical significance and functional form of the temperature terms is only reported in scant detail (P 1139, C1), and significance or otherwise of the humidity effects not reported at all. The use of both dewpoint and daily mean temperature in the analyses may have implicitly allowed some assessment of humidity effects, and Figure 6 in the Schwartz paper (P 1144) shows results for dewpoint temperature. EPA finds that weather factors were adequately evaluated by Schwartz (1993), and were apparently some of the most powerful methods used that were overlooked or ignored by Davis, et al.

Davis et al. state "... if humidity is included -- for which there is strong, though not overwhelming, evidence -- then the PM_{10} effect is not significant, but if it is omitted, then the evidence for a PM_{10} effect is much stronger." Use of the two-day change in specific humidity as a predictor of mortality in the analyses reported by Davis et al. appears to be justified, but the lack of significance of PM_{10} is more credibly attributed to use of a less predictive PM index, the moving average of the current day and two previous days' PM_{10} (lags 0+1+2). As noted above, Sacks et al. point out in their March 12, 1997 submission: "... a significant effect when the three-

day [PM_{10}] average EXCLUDES the current day, but not when it INCLUDES the current day." This appears to have little effect on the mnsh and mnsh2 coefficients, however, and little effect on the size of the PM_{10} regression coefficient, but increases the statistical significance of the PM_{10} coefficient (R.L. Smith, personal communication, Feb. 8, 1997). This suggests that specific humidity is less likely to be a confounder of the combined effects of PM_{10} and one-day lag maximum temperature effects (tmax1 and tg30_1) than a nearly independent predictor after time trend, PM_{10}, and tmax effects are included in the model. On the whole, the revised results reported by Sacks et al. tend to confirm the findings by Schwartz (1993) of a significant PM_{10} (lag 1+2+3) effect using a different modeling strategy. EPA agrees that findings in different locations or data sets may indeed be sensitive to the choice of a PM metric, but EPA suggests that this is more likely due to other site-specific or study-specific differences (e.g., differences in climate and acclimatization to weather extremes, baseline incidence of susceptible subpopulations such as elderly people with predisposing cardio-pulmonary conditions, and differences in other environmental stressors). The revised NISS findings mainly illustrate the importance of using appropriate lag structure.

Davis, et al. indicate that in their initial analysis, normal linear regression is employed together with a square root transformation. The precise form of this model is not comparable to the Poisson-exponential model used by Schwartz (1993), as now typically employed by most analysts examining PM effects where relatively low count daily mortality data occur.

Davis et al. further state that "By making relatively small changes in the specification of the model, the estimated PM_{10} effect varies considerably." However, some of the variations in model specification that Davis et al. apparently consider "small" are in fact among the major issues in model specification, as defined in the PM CD. These include: (1) a sufficient degree of smoothing to remove long-term trends and local "events" of several weeks' duration, probably requiring greater resolution than can be provided by a 21-knot spline over 1250 days of data; (2) a sufficiently flexible model for removing effects of weather on mortality, probably requiring greater resolution than can be provided by linear models for specific humidity and U-shaped (piecewise linear) models for temperature; (3) correctly specified PM_{10} metrics that include possible longer-term effects, which acknowledges that PM effects and weather effects in the Alabama piedmont may indeed differ from PM effects and weather effects in other locations, such as the maritime climates of London or Philadelphia.

Davis et al. also state "... what does make a big difference is deciding which covariates to put into the model. In the case of the Birmingham data, these are of two types, (i) a trend component representing seasonal or long-term effects, and (ii) meteorological variables." EPA agrees that these are important, but has serious reservations about the way in which these have been handled by Davis et al. It is likely that the time tends have not been adequately modeled, since each spline segment covers ca. 60 days of data by a single cubic function. This may not be adequate to capture transient events of one or two weeks' duration, such as epidemics of infectious diseases. The essence of the daily time series epidemiology method is that each given day's health endpoint is related to the levels of environmental stressors (weather, air pollution) over

the preceding few days, at most 5 or 6 days. It is therefore necessary to be reasonably sure that the long-term detrending removes effects down to about this scale. EPA has also expressed some concern about the adequacy of Fourier series used by Schwartz, but Schwartz's sensitivity analyses using the nonparametric STL method showed little sensitivity for the estimated PM_{10} effect, roughly the same finding as Davis et al.

EPA's assessment is that while Davis et al. used appropriate meteorological variables, they did not use sufficiently flexible models to capture known or suspected nonlinear effects of these variables. Preliminary assessment of collinearity among meteorological variables, and of potential covariate confounding among meteorological variables and PM_{10}, might have simplified the model search procedures. Finally, the authors omitted what may be the most important criterion of all: use of a PM_{10} metric that accurately captures the temporal dependence of health endpoint on PM_{10} exposure. EPA concludes that use of pmmean3 (lag 0+1+2), or any other moving average that includes current-day PM_{10}, is inappropriate for Birmingham.

B. Review of Unpublished Report by Roth and Li (1997) on Analysis of the Association Between Air Pollutants with Mortality and Hospital Admissions in Birmingham, Alabama: 1986-1990

Roth and Li (1997) also examined Birmingham mortality data, as well as hospital admissions data; and they also report finding no significant PM associations if they accounted for any non-linear temperature effects that may occur in Birmingham. Roth and Li also found that the choice of pollutant lag influenced the results greatly.

Preliminary findings for the Roth and Li (1997) analysis were presented by Dr. Roth at the May, 1996 CASAC meeting. Epidemiologists and statisticians at the meeting pointed out a number of shortcomings in the analytical strategy and in details of the models being evaluated. The materials from Drs. Roth and Li recently provided to EPA as attachments to public comments show that the deficiencies pointed out at the CASAC meeting 10 months ago have not been adequately addressed. The first factor is that most the models evaluated were unrealistic to begin with. A reasonable model search and evaluation strategy would have rejected these models and simple common-sense arguments would have been sufficient to exclude them from serious consideration. The comments of some CASAC members are particularly illuminating. We quote directly from the transcript for May 17, 1996:

> Dr. Kinley Larntz: "What is very important ... is to actually analyze the experiment he [Dr. Roth] did and to see what the effects are. Some of the negative effects come from zero lags where you expect nothing to have occurred, so it is a negative effect; it should be noise, but it is a negative effect. When you mix those zero lag negative effects with positive effects at other lags, you might get a distribution like he had. ... I think the implication from that presentation was that you get whatever you would like from just

messing around. I think that is not correct. Those analyses are not all independent." [Transcript for May 17, 1996, P 37 L 20 to P 38 L 11]

Dr. Jan Stolwijk: "What he did was to do a number of analyses, most of which somebody who was actually doing a study trying to find out about things would not do. It would predictably give us non-response. If you take the experimental design space ... and that experimental design space includes a large number of things that are not expected to produce a result, you are going to dilute the results that you seek. If you select out of his [Dr. Roth's] matrix the things that other people have done, he comes to a different conclusion than when he takes his whole matrix. ... you are going to get a random effect that shows that there is no effect. He [Dr. Roth] did this, I think, on purpose in this case. Most epidemiologists, I think, have been trained to limit their observations to something that they can state or would have stated before they started and observe that and base their conclusions on it.

"What he did was to do a series of experiments essentially or a series of analyses which most people who are knowledgeable in the area would say would not produce any result, we would not expect to see any result, and he [Dr. Roth] did not see any result. As a result, results that you would have expected to have seen did not stick out any more as notable findings." [U.S. EPA 1996(c), P 45 L 15 to P 46 L 16]

EPA agrees with these CASAC comments. Several points are worth emphasizing.

(1) The 2,400 mortality and 12,000 morbidity computer runs fitted to Birmingham data by Roth and Li did not appear to follow any model specification search method or optimization scheme commonly used in statistical modeling.

(2) A second factor in the significantly small number of significant PM_{10} effects is that the use of barometric pressure as a covariate introduced a potentially significant covariate confounder between PM_{10} and other weather-related variables (temperature and humidity) that have been used by almost all investigators. This is very likely to inflate the estimated standard errors of the coefficients for PM_{10} and for the genuinely predictive weather variables. Since the authors provide no information on the size of the other effects, one must presume that statistical significance outcomes of the study have been badly distorted by inclusion of a covariate confounder; and

(3) A third factor in the remarkable attenuation of statistical significance to less-than-chance significance may be in the misspecification of temporal detrending and weather models, discussed below.

<u>Pollutant Concentrations from Varying Time Periods Used</u>

The exact periods of time included in various R&L analyses for Birmingham are not clear. The title states "... 1986-1990"; but the abstract states: "... The period of the mortality study was from 1988 to 1993; the period of the morbidity study was from 1986 to 1990. In the mortality study, PM_{10} and CO were considered in all years of the study; O_3 in all springs, summers, and autumns of the study; and SO_2 in 1990-1993. In the morbidity study, TSP was considered in 1986-1987; PM_{10} from 1987 to 1990; SO_2 in 1986, 1987, and 1990; O_3 in all springs, summers, and autumns of the study; and CO in all years of the study." Further, Roth and Li state: "The pollutant data by year used in the study were: TSP in 1986 and 1987; PM_{10} from 1987 to 1993; O_3 for March through November from 1986 to 1993; SO_2 for 1986, 1987, and 1990 to 1993; and CO from 1986 to 1990." Thus, the periods of time covered by different Roth and Li analyses are very confusing to sort out and represents a patchwork of dates, depending on data availability, and the reader is given no information about which multiple pollutants analyses covered which periods of time. The mortality studies are said to cover the period 1988-1993. PM_{10} data were available for the later part of 1987 through 1993, representing the longest time series including PM_{10} that could have been evaluated using these data.

O_3 data were available for the months March-November from 1986-1993. Even if O_3 levels were low in the months December-February, there may still have been some O_3 health effects during winter than were attributed to other pollutants. The reader can only presume that in any models including O_3 effects, the winter seasonal component may have been biased (and therefore biased combined pollutant effects estimates for other seasons, to some extent) by omitting the O_3 term.

SO_2 data were available for 1986-1987 (when PM_{10} data were not available) and 1990-1993, but not for 1988-1989 (when PM_{10} data were available). Therefore, analyses of contiguous time series including both PM_{10} and SO_2 could only have been carried out for the years 1990 to 1993, a much shorter period than those analyses including only PM_{10} (latter part of 1987 to 1993). Is there any basis for assuming that PM_{10} and SO_2 exposure and effects were the same over the whole interval? Was this tested and not reported? PM regression coefficients based on a single year of data are likely to be highly variable, since studies with less than 600-800 days of data (two years, roughly) are likely to have little power to detect real effects (CD, Section 12.6.2.1).

CO data were only available for 1986 to 1990. Therefore, it would not be possible to carry out multiple pollutant PM_{10} and CO analyses for 1986-early 1987, since PM_{10} data were not available, nor PM_{10} and CO analyses for 1991 to 1993 when CO data were not available. Only the period late 1987 to 1990 could be studied for PM_{10} and CO. Joint analyses of PM_{10}, SO_2, and CO would therefore have to be limited to only the later part of 1987 and 1990. This is illustrated below in Figure C-1.

Given the very different time intervals that could be evaluated from the available air pollution data, one must conclude that any reported comparisons of coefficients across different multiple pollutant models are likely to be highly misleading (e.g., sample sizes, pollutants, and years included in a given model run could be very different from those in other runs), and would require many more careful comparisons than are shown in this paper.

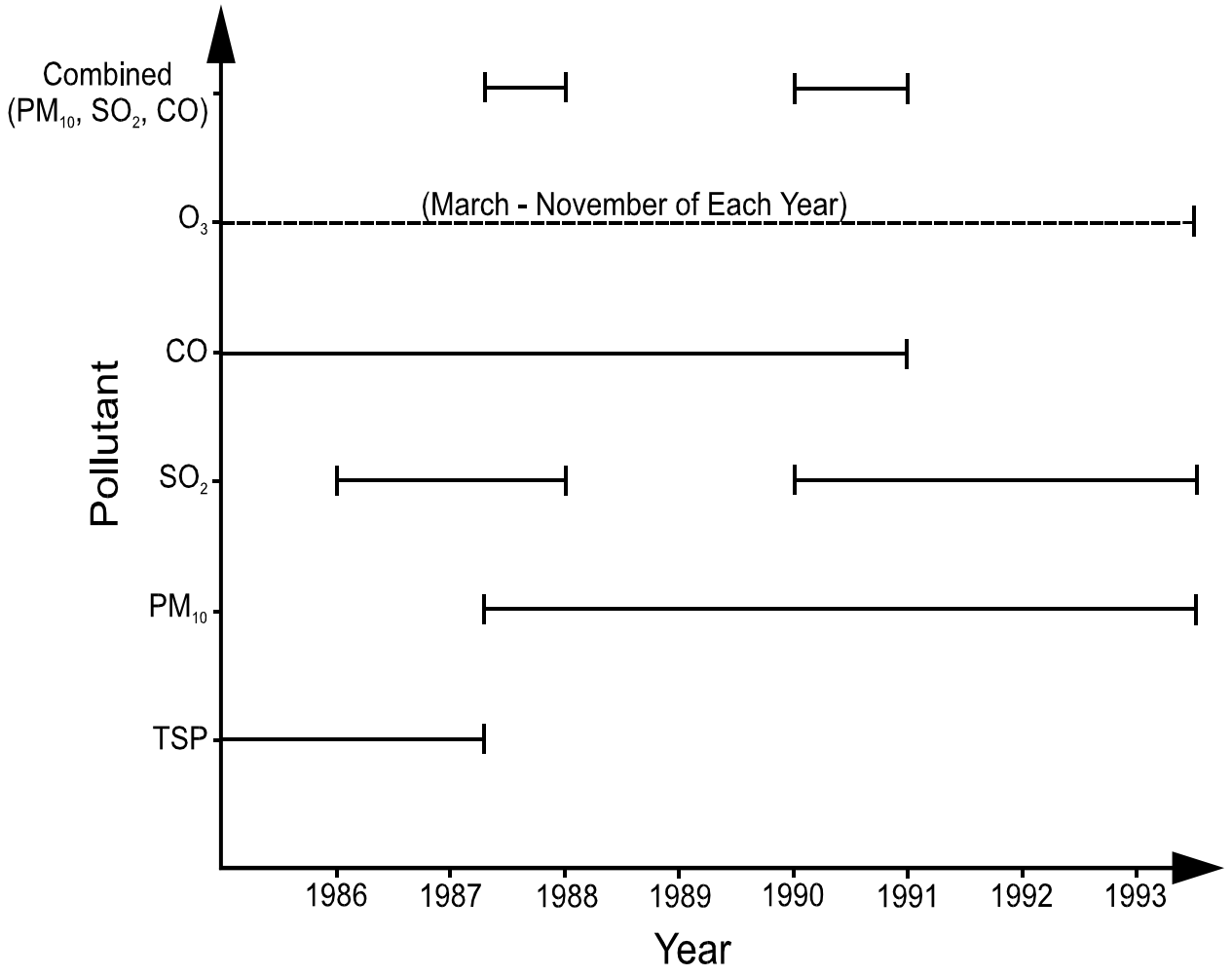


Figure C-1. Birmingham Air Pollution Time Frames used in Li and Roth Analyses

Model Specification Issues

EPA found numerous deficiencies in the modeling strategy used by Roth and Li. Three of the deficiencies are particularly of importance within the present context of response to comments: (1) the analytical strategy used by Roth and Li included a large number of implausible models that would have been rejected *a priori*, or eliminated by standard model specification or optimization search strategies; (2) the V-shaped temperature model used by Roth and Li differs substantially from the U-shaped models found in many other studies by use of nonparametric smoothers, and is thus an implausible alternative that may introduce spurious confounding (as noted during the CASAC meeting); and (3) use of same day (zero-lag) PM-mortality regressions in a number of the models tested by Roth and Li represents a non-credible lag structure set aside

by most knowledgeable analysts as implausible. Interestingly, several of the Roth and Li model runs with PM_{10} lags similar to those proposed by Schwartz, with a PM moving average over the three days preceding the response day but not the current day's PM, show significant effects even in the face of somewhat implausible weather models.

Roth and Li observed that "The Poisson model has been used by many other investigators of PM epidemiological studies, but the deviations from moving averages model might be preferable because it automatically adjusts for time and seasonal trends. We could not find any preference in the statistical literature for one of these methods over the other." This is discussed in great detail in the PM CD (Ch. 12). When the number of daily events is very large (as in the London and New York City analyses cited in the 1986 PM CD, and possibly in the Philadelphia analyses), then the counting data may be adequately approximated as having a normal distribution. Even with Birmingham data in this report, and especially when the data are subdivided by age and by cause or by ICD-9 category, then the expected daily counts are so small that the normal approximation is inadequate and it is probably necessary to use methods based on Poisson or hyper-Poisson models. Appropriate methods for longitudinal data analysis for discrete data were discussed in detail in many reports cited in Ch. 12 of the PM CD and were further developed by HEI (Samet et al., 1995, 1996ab); and one must assume that Roth and Li had some knowledge of these methods, but chose not to use them.

Roth and Li ignore the fact that most analysts now use methods that combine a Poisson or hyper-Poisson statistical model for discrete data with various methods for empirical smoothing, of which deviations from moving averages constitute a particularly limited example. The HEI studies present many adequate applications.

Roth and Li characterize their Poisson regression model as:

$$\text{Log}\,[E(Y_i)] = X_i\,\beta + V_i$$

where X_i is the matrix of covariates on day "I", Y_i is the mortality or hospital admissions on day "I", and

$$V_i = E_i - \alpha_1 V_{i-1} - \ldots - \alpha_p V_{i-p}$$

is the error term and is assumed to be an autoregressive process of a given order p. The estimation method is the Yule-Walker estimation (Gallant and Goebel, 1976)."

This is not a standard specification. The current standard methods (Samet et al., 1995, 1996ab) involve fitting an optimal model for weather and temporal detrending by the use of non-parametric smoothing functions. It is clear that the autoregressive function defined by Roth and Li is a smoothing function that gives zero weight to later values Y(t+x) and is a weighted linear combination p previous residuals (undefined). The time series structure for Poisson process models used by Liang and Zeger (1986) is more appropriate in this context.

The only regression diagnostic provided by Roth and Li is the coefficient of determination, R-squared (R^2). R^2 is not an adequate index for Poisson regression models, since much of the variation in Poisson models with small expectation E{Y(t)} is intrinsic Poisson variability that cannot be explained by any regression model, no matter how perfect. In a Poisson model, the variance of Y(t) (square of the standard deviation) is equal to the mean, E{Y(t)}. In a hyper-Poisson distribution, the variance is greater than the mean. One useful indicator of the goodness of fit of the regression model is to estimate the overdispersion index, defined as (variance/mean). When this is close to 1, the model is adequate. Models developed by Samet et al. for Philadelphia have reduced the overdispersion index to 1.05 or less. There does not seem to be any logical role for a random effects mixed model of the sort proposed by Roth and Li.

The Poisson model with a moving-average smoothing function would be more appropriate than an auto-regressive random process. A simple 15-day moving average would have s = 7 and equal weights 1/15 on each day. Shumway's 19-day filter (s = 9) puts higher weights on the central part of the data and lower weights towards t+/-9. These may be less effective than the use of more recent nonparametric smoothers. Spline functions tend to allow much better local fitting of the data, with each cubic segment nearly independent of all others except at the join points or knots. Often, a larger value of s such 15-45 days provides nearly as good detrending as the shorter periods (s = 7 to 9) with far fewer smoothing parameters, thus produces a larger Akaike Information Criterion (AIC). The autoregressive model assumed by Roth and Li uses only 6 parameters, but imposes an identical autoregressive structure over the entire span of the model. This is likely to provide an extremely inadequate degree of smoothing, and this mis-specification thereby also contributes to reducing the significance of the pollution and weather terms. However, since Roth and Li provide no appropriate information by which to evaluate model fit, one cannot quantify the effects of their smoothing algorithms.

In addition, the authors noted that the predictive power of PM_{10} varied from year to year. Any other result would be surprising. The power of PM studies to detect effects at current levels of exposure requires, as a reasonable minimum, at least 600 to 800 days of data in order to have a marginally acceptable chance of detecting a real effect (PM CD, Ch. 12). Much shorter time series at much higher PM exposures, such as in London winters in the 1960s, produced statistically significant effects for most years (PM CD, 1986; Schwartz and Marcus, 1990), but a 50% increase in mortality in a population of 8 million on a high-pollution day is a high price to pay for greater certainty. Any analyses that use only 365 days of data are likely to have highly unstable regression coefficients, with large standard errors and low significance levels. Roth and Li appear to have verified what is already well known, that many years of data are needed for reliable estimation of contemporary ambient PM effects.

In any case, simple Pearson correlation coefficients calculated on an annual basis may obscure the true seasonal differences in correlation structure. EPA reanalyses of the Health Effects Institute (HEI) Phase IB analyses of Philadelphia data (Samet et al., 1996b) were reported prior to the May, 1996 CASAC meeting and were available for review by Roth and Li. These clearly showed important seasonal differences in correlation structure among copollutants, and

suggested that similar seasonal differences in correlations might exist involving meteorological variables. The report by Roth and Li neglects to evaluate these obvious differences and to take them into account in their analysis.

Model Specification Issues: Weather Variables

The authors stated that the methods that we used to control for temperature were: linear adjustments; deviations from ideal temperatures; and dummy variables. EPA has concerns about all of these methods, even though some of them were earlier used by some investigators. Linear models, even with season-specific regression coefficients, may underestimate the relatively steeper responses at very low and at very high temperatures, and may over-estimate responses during a mid-range of relatively comfortable temperatures and humidity. The use of deviations from an ideal temperature produces a "V-shaped" response function that is likely to have both kinds of problems, underestimating responses at extremes and overestimating responses near the so-called "ideal" temperature. Either kind of model mis-specification may bias the effects attributed to weather vs. air pollution. Some technical issues are discussed in more detail below.

Roth and Li state that they included barometric pressure in their models. This is interesting because barometric pressure is rarely included in daily time series analyses because it is often correlated with other meteorological variables, such as temperature and humidity. In Table C-1, below, the Pearson correlations among the variables analyzed by Roth and Li for Birmingham demonstrate the potential strength of barometric pressure as a "covariate confounder" . Table C-2 focuses more specifically on PM_{10} and three meteorological variables used in the analyses: MXTP (maximum temperature during the day), HUMID (relative humidity), and PRES (barometric pressure). One of the correlations, MAXTP and PRES, appears to be misreported in Table C-1, recopied from R&L. No correlation coefficient smaller in absolute value than 0.064 is statistically significant. It is likely the decimal point is misplaced, and the correct PRES vs MAXTP correlation is 0.336, not 0.0336. In that case, the correlations of MAXTP and HUMID vs. PRES (0.336 and -0.260 respectively) have a very similar pattern to the correlations of MAXTP and PRES vs PM_{10} (0.373 and -0.378 respectively). That is, PRES has a similar linear relationship to MAXTP and HUMID as does PM_{10}. Thus, including PRES in a regression model along with MAXTP and HUMID will induce a collinearity with PM_{10}, thereby inflating the variance of parameter effect size estimates. This will greatly reduce the statistical significance of any possible PM_{10} associations with mortality and morbidity.

.

TABLE C-1. PEARSON CORRELATION COEFFICIENTS IN BIRMINGHAM, ALABAMA							
	RXTP	**HUMID**	**PRES**	**PM_{10}**	**SO_2**	**O_3**	**CO**
Hospital admission	-0.207	0.069	0.046	-0.046	-0.025	-0.030	0.028
Death	-0.076	-0.002	0.023	-0.002	0.012	0.017	-0.018
Matp		-0.058	0.0336	0.373	-0.204	0.44	-0.009
Humid			-0.260	-0.378	-0.217	-0.33	-0.392
Pres				0.116	0.189	-0.15	0.17
PM_{10}					0.197	0.273	0.655
SO_2						-0.064	0.351
O_3							-0.105

TABLE C-2. CORRELATION COEFFICIENTS FROM TABLE RL1		
VARIABLE	PRES	PM_{10}
MAXTP	0.336*	0.373*
HUMID	-0.260*	-0.378*

* = $p < 0.01$

Another area of concern is the specification of the functional form of the temperature effects model. While overall relationships between excess mortality and temperature are U-shaped over the year, these effects are experienced differently by season. A winter temperature-mortality relationship is likely to be monotonic, possibly flat over a range of comfortable ambient temperatures, but increasing with decreasing temperatures below some low-temperature change point or "threshold". A summer temperature-mortality relationship is also likely to be monotonic, possibly flat over a range of comfortable ambient temperatures, but increasing with increasing temperatures above some high-temperature change point or "threshold". The winter low-temperature threshold is much lower than the summer high-temperature threshold. Furthermore, it is highly unlikely that the winter effect per degree below the low-temperature threshold is the same as the summer effect per degree above the high-temperature threshold.

The V-shaped temperature effects model proposed by Li and Roth is among the least plausible of all of the models that have been proposed by various investigators. Roth and Li used deviations from an "ideal temperature" which they took to be approximately 70º. EPA finds this approach to be odd. There is no indication about how this ideal temperature was selected. It is not unlikely that the ideal temperature could vary from season to season, or even from year to year, depending on the incidence and use of climate-control equipment and weather acclimatization among the most highly affected subpopulations. An ideal range of temperatures, say 60-80 degrees F for example, would be more appropriate and better reflect temperature stress study results showing increased heat-stress effects as temperatures exceed 90-95 degrees F and cold-stress effects as temperatures approach or go below freezing (32 degrees F). Even if the "ideal" temperature had also been an adjustable parameter that had been estimated from the data, the difference between the Li and Roth V-shaped function and a U-shaped function of temperature would include a substantial overestimation of the effect of temperature at temperatures somewhat above and somewhat below the "ideal" temperature. This is likely to introduce spurious confounding between weather and air pollution. Other methods, such as using indicator variables for temperature ranges or using nonparametric smoothing techniques, can provide more satisfactory models for weather effects than this V-shaped parametric model.

The use of dummy variables is better than either of these alternatives, since it may characterize non-linear relationships, but still may not be adequate. It is likely that most of the cold-related deaths occur in the lowest decile of days, less than 52 degrees F, and that the response to low temperature varies significantly within that lowest decile. Likewise, it is likely that most of the heat-related deaths occur in the highest decile of days, at least 93 degrees F, and that the response to high temperature varies significantly within that highest decile. It is likely that a model allowing for variations within each decile would be more realistic.

Roth and Li explained that, considering that temperature and dewpoint were highly correlated, weather variables used in their analyses were maximum temperature, relative humidity, and barometric pressure. Obviously, other choices could have been made. As noted above, barometric pressure is a covariate confounder of PM_{10}, maximum temperature, and relative humidity. Was any preliminary assessment made of the predictiveness of these variables relative

to other indicators, for example: maximum temperature in summer, minimum temperature in winter; mean temperature; specific humidity etc.? What about interactions, such as Schwartz's 'hot humid days' variable? Was the predictiveness of any model with only weather evaluated? What was the model-building strategy used, if any?

Model Specification Issues: Co-pollutants

Roth and Li note that there was little difference in the significance of PM_{10} effects when PM_{10} was analyzed alone or in combination with other pollutants. Given that the Roth and Li analyses with PM_{10} alone had little power to detect effects (due to structural deficiencies in the methods used) then it is not surprising that, when correlated copollutants were added to the models, the power of the analyses was further diluted and the results were even more significantly non-significant. A number of analytical approaches were suggested in Ch. 12 of the PM CD that could be used to evaluate the effects of correlations among copollutants and weather variables. Due to significant data gaps for air pollution in Birmingham, admittedly not all of these analytical approaches could have been used to evaluate copollutant effects; however, Roth and Li used none.

Roth and Li further state that, for complex reasons, incorporating additional pollutants into the analyses has a major impact on results in some cities (Li and Roth, 1995 on Philadelphia) whereas in other instances (such as Birmingham), it has at most a minor effect. This is due to correlations between different pollutants, weather effects, day of week effects, etc. While these analyses can be complicated, EPA notes that other investigators have made considerable progress in pursuing detailed analyses to informative conclusions. These include substantial separation of PM effects from weather effects (Samet et al., 1996ab; Pope and Kalkstein, 1996), and separation of some copollutant effects (CO and O_3) from others (TSP, SO_2, NO_2 in Philadelphia, Samet et al., 1996b, discussed further in the PM CD). While it may not be easy to detect separate copollutant effects in all studies, there are a number of studies carried out in different locations in which it does appear possible to identify a separate PM effect that exists over and above the combined effects of a mixture of urban or regional air pollutants. EPA regards this problem as complex and difficult, requiring careful attention to data and to methods, but resolvable with current advanced techniques and data sets.

With regard to lags used in their analyses, Roth and Li state that the lag structures of pollution that they considered are: the average of the current and previous days, the current and the two previous days, and the average of the previous three days. The significance of pollutant lag days varies from city to city." EPA notes that the earlier findings by Schwartz (1993, 1994b) suggest that lag 0 (current day) effects have little or no correlation with health endpoints, so all models including this term are *a priori* less plausible, and models including some or all of days 1, 2, and 3 are more plausible for predicting mortality effects. Also, as noted above, the inclusion of barometric pressure as a covariate confounder is likely to substantially attenuate the statistical significance of PM_{10} and weather as predictors of mortality.

Some of the most important models needed to evaluate these results are not presented by R&L -- those models with no air pollutants. A basis for comparison could be provided by models with temporal detrending and with evaluation of the adequacy of weather-related terms. Finally, all of the models evaluated are simple moving averages or one-day exposures. Models with weighted moving averages may also have proven more informative.

Based on their thousands of model runs, Roth and Li conclude that the Birmingham data on mortality and hospital admissions show little indication of a PM (or TSP) effect. In about half of the analyses, PM_{10} (or TSP) was positively associated with mortality or hospital admissions, and in about half it was negatively associated. But the overwhelming majority of these results were insignificant. Several factors built into the analyses virtually guaranteed that this would occur. The most important factors are: (i) most of the models evaluated were known to be implausible and would have been rejected *a priori* or would have been rejected by any reasonable model specification search strategy; (ii) inclusion of barometric pressure, a covariate confounder of the association between PM_{10}, temperature, and humidity, may have inflated the standard errors and thereby deflated the statistical significance of all of the linearly related terms; (iii) the weather models were all more or less misspecified; (iv) no effort was made to separate multiple pollutant models, accounting for or evaluating pollutant correlations season-by-season, over periods of time that may have had little overlap and were much shorter than the whole study period; (v) the statistical model may have been misspecified, especially the Poisson autoregressive regression model; (vi) analyses by year and season may not have adequately detrended the time series, and (vii) analyses using one-year subsets of the data had very little power to detect PM effects. The results showed far fewer "statistically significant" findings than would be expected purely by chance.

Roth and Li state: "A critical question to be raised about our study is which of all of the analyses performed should be given highest priority. It is apparent that other investigators have also grappled with this issue because they rely on different models to analyze the data in different studies."

In fact, Roth and Li have produced an astonishing result. The authors found only 19 significant PM_{10} hospital admission results out of 12,000 analyses conducted (0.16%), and 7 significant associations with mortality out of 2,400 analyses (0.3%). One would expect that even if there were no real PM_{10} effect, PM_{10} would be identified as statistically significant at the $p< 0.05$ level about 5% of the time. In other words, one would expect to find 120 PM_{10} analyses that meet the criterion for statistically significant coefficients out of 2,400 analyses. While the analyses are not completely independent, differences in pollutant lag structure, weather model, and statistical model are likely to provide a significant degree of independence, so that the finding of 7 significant effects in the mortality analyses, compared to 120 expected, is significantly **smaller** than would be expected purely by chance. This again suggests that several factors may have been built into the analyses that attenuated statistical significance to even less-than-chance levels.

The critical question in the study by Roth and Li is whether *any* of the 2,400 mortality models or 12,000 hospital admissions models fits the data well enough to be taken seriously. The authors report only R-squared statistics, which are known to be inadequate for characterizing goodness-of-fit of Poisson regression models with small expected counts, and the R-squared statistics are cited as being in the supporting documentation, not in the report itself. They provide no statistics appropriate to Poisson regression, such as deviances, overdispersion indices, or Poisson AIC criteria. There are no visual or graphic diagnostics. This report is virtually unique in its apparent indifference as to how well their models describe the data at hand.

Almost every other study reviewed by EPA has stated a more or less explicit model search or model specification strategy, with appropriate criteria for evaluating different models. Some form of "optimization" is needed because not all of the parameters of these models are known in advance, nor is the exact functional specification. This introduces some conceptual issues, which are discussed in detail in Chapter 12 of the PM CD. Section 12.6 offers many examples comparing results from different models. The use of formal criteria such as AIC for comparing non-nested models with different specifications is illustrated well in studies such as Samet et al. (1996a,b). The statistical community generally accepts the necessity of comparing only models that are "best of the class", and rejecting models that do not satisfy some sort of plausibility criteria or goodness-of-fit criteria, as pointed out by CASAC statisticians and epidemiologists. Roth and Li provide no basis for answering the question they have posed, whereas other investigators have no difficulty in proposing criteria for model acceptability. One may disagree with the criteria, or with their specific applications, but few, if any, other investigators besides Roth and Li have fitted models in the virtual absence of specific criteria for when a model is "best", or at least "good enough".

In summary, Roth and Li clearly conducted extremely inappropriate reanalyses of Birmingham mortality and morbidity. Many of the approaches used by Roth and Li are unreasonable and implausible. These include:

(1) Use of different ranges of years for copollutant models;
(2) Inclusion of current-day PM_{10} in mortality models, when most previous studies for Birmingham find the current-day terms nonsignificant or weakly negative;
(3) Use of dummy variables for season and year, rather than long-term smoothing functions;
(4) Use of linear models or absolute deviation models for weather;
(5) Use of dummy variables for weather deciles without evaluating inclusion of a parametric model for extremes of heat and cold;
(6) Use of barometric pressure, a potential covariate confounder for PM_{10}, temperature, and humidity;
(7) Use of an autoregressive model with 6-day lags for smoothing the mean;
(8) Apparent lack of use of properties of the Poisson distribution, and lack of use of model diagnostics appropriate to the Poisson distribution.

Since no diagnostics of model adequacy are presented, it is possible that none of the 14,400 models used by Roth and Li are adequate. There is neither an empirical basis nor a theoretical basis for regarding these models as demonstrating "reasonable approaches". It is clear that their very incongruous model choices cannot be accepted over other more valid methodological approaches used in most other studies of PM effects — several of which find statistically significant effects of PM on mortality and morbidity in Birmingham, using more appropriate lags and weather variables.

C. Review of S.H. Moolgavkar, E.G. Luebeck, and E.L. Anderson (1997) Air Pollution and Hospital Admissions for Respiratory Causes in Minneapolis-St. Paul and Birmingham, Epidemiology (1997) 8:364-370

In this analysis, Moolgavkar, et al. investigated the association between air pollution and hospital admissions for chronic obstructive pulmonary disease (COPD) and pneumonia among the elderly in Minneapolis-St. Paul and Birmingham, Alabama, over the five year period between 1986-1991. These associations were investigated using a generalized additive regression model that controlled for temperature, day-of-week, season and temporal trends. Pollutants included in these analyses were PM_{10}, SO_2, NO_2, O_3 and CO in Minneapolis-St. Paul and PM_{10}, O_3 and CO in Birmingham. To evaluate the effects of multiple pollutants on the regression coefficients these pollutants were included both individually and simultaneously in the analyses. When considered individually, the authors found positive and significant associations for all pollutants, except CO, with total respiratory hospital admission (COPD + pneumonia) in Minneapolis-St. Paul, of which ozone showed the most strongly significant associations. Ozone remained significant with the inclusion of all four significant pollutants in the model. Similar pollution/health effect associations were not demonstrated in Birmingham, Alabama. From these analyses, the authors conclude that there is an association between air pollution and respiratory hospital admissions in Minneapolis and of the pollutants considered, ozone was the most strongly associated pollutant. Moolgavkar et al. go on to say that while PM_{10}, SO_2, and NO_2 demonstrated positive associations, it is not possible to single out one pollutant as being more important than the others.

Several germane questions provide a useful framework for evaluating at least some key aspects of the Moolgavkar study:

(a) Is the general method of analysis of the Birmingham data sound and appropriate?

(b) Has the study appropriately handled pollutant and weather lag structures and detrending for control of temporal variations in the data?

(c) Are the conclusions of no PM effect in Birmingham consistent with the data? What factors might account for different findings than Schwartz and Dockery (1993)?

As for the first point, the overall approach used by Moolgavkar is consistent with the general approach taken by most air pollution epidemiologists, in that they have attempted to address

several key issues of concern. These issues include temporal trends, variations in hospital admissions due to weather, the effects of day-of-week and pollutant lag structure. Adequate control of these potential sources of confounding is very important in assessing the effects of pollution on human health. The authors chose to control for seasonal cycles in the data using non-parametric smoothing techniques. The Akaike Information Criterion (AIC), a statistic which assess the overall fit of a model while including a penalty for the number of parameters in the model, is used an "exploratory tool" in helping with model choice. The authors state that the AIC should not be used to choose a model when other information is available, but in the absence of biology the AIC can be used to determine the "optimal degree of smoothing." With this, the authors found that the AIC decreased to a minimum value with increased smoothing, and thus opted to use the smoothing spline that produced the lowest AIC.

The authors are correct that the AIC is to be used as an "exploratory tool" but do not use it as such, but rather as a model-making decision tool. The purpose of smoothing the hospital admissions data is to model the seasonal cycles, not necessarily to get the model which best fits the data. This may or may not be the same point at which the AIC is minimized. In the case of seasonal cycles, other information is available to the investigator. For example, a visual inspection of the actual plots of the outcome series are useful in observing if the seasonal cycles are properly fit, or if more detail than the seasonal cycles (i.e., other phenomena, possibly including pollution associated fluctuations) are being fit. This type of inspection is especially important in the case of hospital admissions data, where long-wave cycles are not as dramatic as in typical mortality series. Inspection of plots of the residuals versus fitted values are also an important source of information which can be used to guide the investigator to final model specification decisions. Furthermore, the use of the AIC (or any other goodness-of-fit criteria) as the sole determinant of the extent of long-wave cycle control fails to take into account potential autocorrelation or overdispersion in the outcome series, both of which can influence the pollutant coefficient size and significance. In the case of Birmingham, the lowest AIC was produced from a smooth with 60 degrees of freedom. This is equivalent to a 37-day moving average, which may not be conservative enough. As a result, the Birmingham hospital admissions data may be autocorrelated. This is a potential source of major confounding. In contrast, a smoother with 130 degrees of freedom was used in the Minneapolis-St. Paul analysis, which is roughly equivalent to a 17-day moving average. This may be too conservative, thus potentially filtering out variation that may not truly be attributable to season, but to short-wave phenomenon, such as acute air pollution effects.

With regard to how appropriately the study handled pollutant and weather lag structures and detrending, the authors first explored the association of 0-3 day lags of the individual pollution variables with respiratory hospital admissions and used the lags which produced the highest correlations, regardless of the magnitude of the AIC. This is an appropriate way to handle the potential delayed effect pollution may have on human health.

The modeling of weather effects is also an important issue, as mentioned above. The authors control for temperature effects by using a smoothing spline of either a 7-day running average or the mean same day temperature. The use of the 7-day running average is not an appropriate method of controlling for the acute affects of temperature on human health, as shorter-term

effects will not be fully accounted for in such a specification. Also, the authors did not consider any temperature-relative humidity interactions, which could be important, especially in Birmingham. While they state that a strong association between daily temperature and respiratory hospital admissions persists up to 10 days, the strongest associations are most likely observed during the 0-3 day lags after initial exposure. Correlations for days beyond this are likely to be smaller, and may even be in the opposite direction as a result of the common 3-4 day cycle in weather systems (and, hence, in pollution). Thus, averaging these later lags in with earlier lags might affect the estimated acute effects of temperature on respiratory hospital admissions, which may, in turn, influence the pollution effect estimates up or down, depending on their interrelations. This point is demonstrated in their Figure 3 where the relative risks for PM_{10} are diminished when the 7-day running average of temperature is used. Less variation in the data was accounted for with the 7-day running average, causing the pollutant coefficients to become biased. It would have been more productive to first investigate the strength of these weather associations for the individually lagged days (as done with the pollutants) and determine the most significantly associated lag (or lags) and include those individual days.

Detrending was conducted using a non-parametric spline smooth of day of study on respiratory hospital admissions. As discussed above, the degree of smoothing was based solely on minimizing the AIC. The use of non-parametric smoothing techniques has become popular in time-series investigations, and has been shown to be an effective method of controlling for temporal variations in the data (such as long-wave cycles and trends), if used appropriately. Such methods should be used with the same *a priori* information and diagnostic evaluations as traditional least squares regression approaches. Inappropriate model specification can lead to confounding, and produce biased regression coefficients.

Without reanalyzing the Birmingham dataset independently, it is difficult to assess whether the authors' conclusion of no PM effect is consistent with the data. However, the modeling specification differences between the two studies may account for such different conclusions. The Schwartz and Dockery analysis undertook various modeling approaches for temporal cycles, weather effects, and the exclusion of influential observations. The resulting Schwartz and Dockery estimates of PM_{10} and ozone were robust, as consistent effect estimates were obtained. On the other hand, Moolgavkar used one basic approach for controlling long-wave cycles and two approaches for controlling weather effects. The method of controlling for long-wave cycles is the most likely the reason for the differing results between the authors' analysis and the Schwartz and Dockery analysis. The potential misspecification of the long-wave cycles by Moolgavkar may have resulted in autocorrelated errors, which can produce biased regression coefficients. Of the various approaches used by Schwartz and Dockery to control for seasonal variations, one approach utilized a spline with 12 degrees of freedom (3 month time interval). This interval is longer than the one used by Moolgavkar and inevitably results in autocorrelated residuals. However, unlike Moolgavkar, Schwartz and Dockery test for autocorrelation in the residuals. If autocorrelation is found, they employ Liang and Zeger's (1986) generalized estimating equations, which correct for autocorrelated data. Moolgavkar apparently failed to address possible autocorrelation. It is likely that pollution was highest in summer (especially for ozone), while daily hospital admissions counts are generally at a minimum in the summer months. Thus, the uncontrolled long-wave associations in the Moolgavkar analysis would tend

to artificially diminish the pollutant coefficient estimates, consistent with the results reported by Moolgavkar.

Another difference between the two studies is the specification of weather. Schwartz and Dockery use two different approaches: the first utilizes seven categorical indicators of temperature and humidity, the second includes linear and quadratic terms for temperature with a linear term for relative humidity. As discussed previously, Moolgavkar controlled for temperature with two approaches, same day mean temperature and a 7-day average of temperature, but did not address relative humidity at all. Thus, the differing Moolgavkar results obtained by Moolgavkar and his associates may also be in part a consequence of the less complete specification of weather employed in their analyses.

With regard to methods of analysis, the authors state that they used indicator variables for day of week and for season and non-parametric smoothers for temperature and day of study. The basic model used by Moolgavkar et al. is over-parameterized in some ways and under-parameterized in other ways. The purpose of non-parametric smoothers for time trends is to eliminate the effects of long-term secular trends, and shorter-term trends attributable to seasonal changes, flu epidemics, and the like. This allows the time series methodology to look at short-term responses to changes in pollution and weather in the preceding few days. Therefore, the use of seasonal indicators along with non-parametric smoothing seems redundant and even inaccurate, since seasons do not start exactly on December 1, March 1, etc. every year. Additional analyses using only seasonal indicators and simple parametric trends (linear, quadratic) should be compared with those using non-parametric smoothers without seasonal indicators, or the seasonal effect should be more appropriately included as a modifying factor on weather and pollution. For example, it is unlikely that response to a 50 degree F day in February has the same health effect as a 50 degree F day in August. In any case, the weather model may be inadequate without including a humidity variable as well, since many studies have found that humidity has an independent effect, or in some cases an interactive effect with temperature.

The authors also indicate that they examined lags of 0, 1, 2, and 3 days for each of the pollutants, and picked the lag that yielded the strongest association with respiratory admissions. This model selection strategy seems highly inappropriate. The authors have assumed that the bulk of the effects of pollutant exposure could occur on any day. There is no biological basis for this assumption. Alternative models looking at moving averages or distributed lag structures would be more reasonable. For example, in Table 3 on p. 16, Minneapolis pneumonia admissions are attributed either to today's PM_{10}, yesterday's ozone and NO_2, or SO_2 from two days ago. Minneapolis COPD admissions are attributed to today's ozone, yesterday's PM_{10}, or CO from 2 days ago. It is not reasonable to assume that air pollution health effects are so narrowly segregated by day of exposure. A much broader class of exposure durations should be considered, possibly extending to lags 4 and 5.

The authors also investigated the lag structure for temperature and found that a strong association between mean daily temperature and respiratory admissions persisted even with a lag of 10 days. They then controlled for temperature by using a smoother with 10 degrees of freedom either for the current day mean or for a seven-day running average. These specifications

seem to be at odds with findings of other investigators. The authors do not report analyses using other temperature variables, such as winter minimum or summer maximum, nor do they report findings using other moving averages or distributed lag structures for temperature. They do not seem to consider the possibility of different lag structures in different seasons, as in mortality studies where some data suggest 3 or 4 day lags in winter and 0 to 1 day lags in summer. Nor were any assessments using humidity as an independent variable or as a modifying factor for temperature reported.

It is also indicated by the authors that, with all covariates in the model, smoothers were used with varying degrees of freedom for the day of study to control for temporal trends. There may be some confusion about the authors' model specification search strategy. Pages 4 and 5 suggest a forward selection strategy, first including a basic model with day of week indicators, seasonal indicators, nonparametric smoothers of time, and then evaluating temperature effects models. Once the baseline model is determined, either this model is fixed and pollutant models are evaluated, or the parameters of the baseline model are re-estimated with each pollutant model. Once the pollutant model is determined, the temporal smoothing model is re-fitted, a limited form of "forward-backward" search. It is not clear whether the search strategy was specified in advance, or constitutes an ad hoc exploratory procedure. *Ad hoc* procedures may be useful, but the reader should be told which was used.

Based on the above noted methodological approach, the authors report that in Birmingham, both PM_{10} and ozone showed generally inconsistent associations with hospital admissions, depending on the lag chosen. EPA notes that, by not using two- or three-day moving averages (lags 1+2+3 by analogy with Birmingham mortality studies), the power of this Birmingham study to detect air pollution effects is weakened. Since the authors have reported little of their findings for Birmingham, EPA is unable to comment in detail on these results.

The authors also report that temperature was associated with hospital admissions even with lag times on the order of 10 days. In their analyses, they considered either a 7-day running average or the mean temperature on the same day. In Minneapolis, the 7 day running average was more strongly associated, whereas the opposite was true in Birmingham. These findings leave open the possibility that the weather model in this paper has not been adequately specified. There are several possibilities: (1) the wrong temperature variables were used; (2) other relevant weather variables, such as humidity or dewpoint temperature, should have been included in the model; (3) weather effects required seasonal adjustment or modification; (4) there was undetected confounding involving season, day of week, or weather.

The authors further indicated that, in Minneapolis, they found a strong and consistent association of ozone with total daily respiratory admissions. Although PM_{10} was strongly associated with respiratory admissions when considered as the only pollutant in regression models, its effect was greatly attenuated when ozone, NO_2, and SO_2 were simultaneously considered. By comparing figures 1 and 2, it can be seen that ozone continued to be strongly associated with respiratory admissions. Among PM_{10}, NO_2, and SO_2, however, no single pollutant stood out as being more important than the others.

Actually, close inspection of their Figure 1 suggests that including both ozone and PM_{10} only slightly reduced the magnitude of both pollutant effects, each of which remained strongly significant. In their Figure 2, the effects of 1-day lag PM_{10}, ozone, and NO_2 were attenuated by smoothing. Both PM_{10} and NO_2 had positive but non-significant effects when more and more information on temporal variability was smoothed out of the data. SO_2 was not statistically significant in any of the models and had a relatively stable magnitude. Since SO_2 did not have a correlation greater than 0.12 with any pollutant, it is likely that SO_2 could have been dropped with little change in results. The correlation of PM_{10} and NO_2 in Minneapolis was 0.31, PM_{10} and ozone was 0.25, but no information was provided on seasonal differences that may prove important. Longer moving averages would almost certainly have provided more incisive inferences.

D. Review of F.W. Lipfert and R.E. Wyzga (1997) Paper on Air Pollution and Mortality: the Implications of Uncertainties in Regression Modeling and Exposure Measurement, Accepted for Publication in J. Air & Waste Manage. Assoc. 47:517-523

The Lipfert and Wyzga (1977) paper is primarily a discussion and evaluation of the paper by Schwartz, Dockery, and Neas published in 1996 (and referred to as SDN), which generalizes and extends comments on measurement error issues made by these authors in earlier papers (Lipfert and Wyzga, 1995a; Lipfert, 1995ab). SDN is a analysis of daily time series data from the Harvard Six-Cities study in which daily mortality counts are related to daily changes in weather and in several indicators of airborne particle mass concentration (PM). This paper by Lipfert and Wyzga (denoted LW97) asserts that SDN has not adequately analyzed their data, and that SDN has not adequately interpreted the results of their analyses.

The conclusions reached by SDN are that fine particles (FP or $PM_{2.5}$) are generally better predictors of excess mortality than the coarse fraction (CF) of thoracic particles ($PM_{10-2.5}$) or other particle indicators evaluated (PM_{10}, $SO_4^=$), suggesting that FP is a much more probable cause of excess mortality at these six Eastern and Midwestern sites than is CF. LW97 argues that CF is much more poorly measured than FP, thus CF will appear to be a much less significant explanatory variable in regression analyses than FP, even if both were in reality equally strong predictors of excess mortality or other adverse health effects.[2] LW97 cite a number of earlier studies and presentations in support of their claims. LW97 also present theoretical statistical arguments suggesting that even if CF had a strong causal relationship to mortality, the excess mortality might be inappropriately attributed to FP because FP and CF are often correlated and

[2]It is of note that CASAC and staff accounted for this possibility and factored it into both the Staff Paper and CASAC recommendations. The Staff Paper notes: "While greater measurement error for the coarse fraction could depress a potential coarse particle effect, this would not explain the results in Topeka relative to other cities. Even considering relative measurement error, these results provide no clear evidence implicating coarse particles in the reported effects." (U.S. EPA, 1996b; p. V-64).

because FP is much more precisely measured than CF. In other words, even if FP had little or no real effect, the CF effect could be attributed to FP because CF is more poorly measured.

EPA disagrees with the assertions in LW97 that measurement errors and related issues invalidate the SDN findings. The first issue is that while the distortion of PM effect size estimates by instrument or analysis errors is theoretically possible, there is little indication that these errors are large enough or so highly correlated as to make these effects likely. The second issue is that under conditions in which one PM index is indeed so highly correlated with another air pollution index that one can almost be used as a surrogate for the other, then in fact it may be appropriate to use the more precisely measured indicator as the health effects predictor in an epidemiology study. The third issue is that the "exposure measurement" issues discussed at length in LW97 may in fact have a very different type of "measurement error" structure that will cause little or no distortion of effect size estimates in SDN or other epidemiology studies. In recent theoretical evaluations by EPA staff and by non-EPA expert consultant statisticians (summarized in Appendix D), the most likely effects of measurement errors are to attenuate the apparent or estimated effects of both the more precisely measured predictor as well as the less precisely measured predictor. In other words, the FP effects reported by SDN for mortality are likely to underestimate true effects.

In addition to amplifying on the above points, some other points raised by LW97 are also noted and commented on below. Instrumental errors were noted by LW97 as one of the sources of exposure errors in the SDN analysis of effects by particle size. LW97 state that dichotomous sampler (dichot) design that has been in common use measures fine particle mass more precisely than coarse particle mass. They noted that a precision of +/- 5 $\mu g/m^3$ as about the best that could be achieved with dichots [Ref. 9: Spengler et al., 1980]; this amounts to about 28% of the fine particles, on average, but 43% of the coarse particles (52% in Boston, St. Louis, Knoxville, and Portage). When dichotomous samplers were evaluated in side-by-side testing [Ref. 10: Camp, 1980], variabilities of +/- 16% were found for FP and +/- 44% for CF. This difference creates a large bias in favor of fine particles, according to LW97.

EPA notes that other investigators (Rodes, et al, 1985) report smaller instrumental error for dichots than indicated above, i.e.: for FP 3.6 to 11.4% and for coarse particles 2.6 to 18.3%.

The concentration-response functions are largely determined by responses to the higher-level concentrations. $PM_{2.5}$ concentrations for all 6 cities had a 75th percentile value of 23.0 $\mu g/m^3$ and a 95th percentile value of 43.1 $\mu g/m^3$, CF had 75th and 95th percentile values of 15.5 and 30.1 $\mu g/m^3$ respectively, and PM_{10} had 75th and 95th percentiles of 38.0 and 67.8 $\mu g/m^3$. This suggests that the relative accuracy of measurements in the upper range was much better than calculated by LW97, respectively 22% for FP and 32% for CF at the 75th percentiles, less at higher percentiles, assuming +/- 5 $\mu g/m^3$ error for both FP and CF. This is not an excessive amount of variation relative to the population variability, so that the difference in measurement error between CF and FP should not be greatly biased in favor of FP.

The relative measurement error indices are, in any event, relatively useless in providing any realistic assessment of measurement error effects. The relevant index of the size of the

measurement error is the ratio of the standard deviation of the measurement error to the standard deviation of the true predictor (measured without error) in the whole population (Fuller, 1987; Carroll et al., 1995). Relative precision indices are in fact coefficients of variation, which measure the standard deviation of the observation relative to its true mean value, not to the variability of the variable in the population.

Lipfert and Wyzga indicate that the statistical consequences of an imperfect (i.e. biased or imprecise) measure of human exposure or dose to the target organ(s) in an epidemiological study include reduced measures of regression "fit" and biased estimates of slope. All epidemiological studies of community air pollution are based on the use of data from centrally located outdoor monitoring stations as surrogates for individual doses or exposures. The differences between the values obtained from ambient monitors and actual personal exposures are "exposure errors"; the constituent sources of such errors are discussed below.

EPA notes LW97 do not distinguish among different kinds of so-called measurement error that have very different consequences for the analyses and may be addressed by very different kinds of analytical methodology. The first error type is the "exposure error", the second type may be called the "ambient measurement error" and is described below. Exposure error arises when the measured PM value at the central monitoring station, or a composite measurement from several regional stations, is assigned as the daily ambient exposure concentration of the population. In population-based studies, where the response is the hospital admissions count or mortality count in the population, this may be an appropriate example of the kind of measurement error called a *Berkson error model*. This is discussed in basic monographs such as *Measurement Error Models* by W. Fuller (J. Wiley, N.Y., 1987, pp. 82-83), and in *Nonlinear Measurement Error Models* by R. Carroll, D. Ruppert, and L. Stefanski (Chapman and Hall, N.Y. and London, 1995, pp. 3, 9, 52-56, 67, 143, 150-151, 255), with extensive references to earlier papers. The Berkson model arises from the fact that the investigator in the daily time series study value assigns the same PM value to characterize the ambient personal exposure concentration for all individuals in the population, whereas the actual personal exposure to ambient air pollution (some of which may occur indoors by PM that infiltrates the building) is somewhat different for each of the individuals. For air pollutants with a regionally uniform distribution, such as $PM_{2.5}$ or sulfates (and PM_{10} to a lesser extent), for which PM is a reasonably unbiased estimate of average population ambient exposure, the effect on statistical estimation may be minimal. As Carroll et al. (1995, p. 150) point out: "The Berkson additive model has an unusual feature, in that for linear regression the naive analysis ignoring measurement error gives correct inference about the regression ..." It is likely that this condition applies to ambient FP or PM_{10} measured at community monitoring sites as an estimate of average population FP and PM_{10} exposure to particles of outdoor origin.

The other measurement error is that relating the measured PM to the true ambient concentration at the central air monitoring station(s). In this case, the classical measurement error model may be applicable, and some attention may need to be given to estimating the effects of such errors. EPA's assessment is that "ambient concentration measurement errors" are unlikely to have such large magnitude or such strong correlations with weather or with co-pollutants as to induce substantial quantitative modifications of effects size estimates.

Lipfert and Wyzga further added that "In nonstatistical terms, adding random noise to the independent variable 'stretches' its scale ... and thus reduces the slope of the relationship ... the slope also biases the threshold to the left ... errors in exposure measurement can obscure the true relationship of a nonlinear ERF ... In multivariate regressions ... accurate representation is just as important for the confounding factors ... the pollutant with the least error will predominate."

EPA agrees that these statements are generally accurate summaries of well-known mathematical properties of regression models with measurement errors under "standard" conditions. The standard conditions generally included additive, independent, normally distributed errors with common variance. However, almost any of these statements must be appropriately qualified when a more realistic measurement error model is proposed, with both additive and multiplicative errors, unknown and highly skewed error distributions, correlated errors, and variances that depend on mean levels. Sometimes, even with valid assumptions, standard methods for measurement error adjustment may not apply to strongly nonlinear models with large measurement errors (for example, logistic regression with response rates close to 0).

The major concern here is that measurement error has attenuated estimates of the real effects of some poorly measured predictors to the point that they are no longer statistically significant and thus "detectable". Many so-called "negative" studies may be of this type.

Varying extents of exposure measurement errors across pollutants and across cities can be important factors in defining the relative significance of pollutant coefficients, but this is not the only, or even the dominant, factor. For example, of the pollutants considered in this paper, ozone has among the highest outdoor spatial correlations over time, and is known to have few indoor sources with most indoor O_3 due to penetration of outdoor O_3 to the indoors (e.g., typical office O_3 indoor/outdoor ratio =0.6), except in buildings with high percentages of recirculated air (Hayes, 1991). Thus, ozone should, according to LW97's logic, have one of the highest correlations with mortality. This turns out not to be the case, with ozone having among the lowest elasticities reported by Lipfert and Wyzga (0.010), despite being among of the measures with the smallest spatial errors. Thus, differential errors alone cannot explain the differences found in the explanatory power of the various pollutants.

However, it is not clear that this effect, if it were able to be accounted for, would significantly change the conclusions drawn from previous studies (i.e., that adjusting for the errors of measurement of coarse particles would increase this parameter to significance in these analyses). For example, in the Ozkaynak and Thurston (1987) cross-sectional analysis, the mortality t-statistic for FP in the entire dataset (n=98 metropolitan areas) was 2.75 ($p<0.01$), while the t-statistic for IP (i.e., PM<15 μm) was only 1.0 ($p>>0.05$), indicating that the differences between the PM components' spatial variability would have to very large, and (just as important) would have to represent a major percentage of these PM measurements' total day-to-day variance in order to account for these very large t-statistic differences. Similarly, in the SDN study using time-series methods, an overall increase in daily mortality per 10 $\mu g/m^3$ equal to 1.5% (t=7.41) was reported for $PM_{2.5}$, but only a 0.4% increase (t=1.48) was found for 10 $\mu g/m^3$. Again, it would take extraordinarily large relative spatial errors (i.e., many times greater for CF than for $PM_{2.5}$) to account for such large differences in significance between these two PM components.

Lipfert and Wyzga state that: "Next, outdoor air quality varies in time and space in different ways for different pollutants so that monitoring networks ... describing large cities or metropolitan areas may not always have been appropriate. ... may not accurately represent the outdoor air quality... especially for pollutants of local origin."

EPA notes that fine particles generally have a very uniform regional spatial distribution (Burton et al., 1995), while PM_{10} is only somewhat less uniformly distributed. Fine and coarse particles can both substantially infiltrate most housing units, so that spatial variability does not appear to be a significant source of error in characterizing regional population exposures to ambient PM, especially in characterizing effects of day-to-day changes in pollution levels, which is the basis for the time series mortality studies. We assume that in most locations, FP has no local sources (this may not always be true) and is otherwise influenced only by weather, whereas CF has more local sources and may be responsive to changes in wind speed and direction.

The differential spatial-temporal variability among monitoring sites is often complicated. For example, two sites could have the same annual or monthly mean concentrations, but have little or no correlation across time. Conversely, two sites may have near-perfect correlation over time, but the concentration at one site is always half that of the other. This may indeed be a more serious problem for CF than for FP. For example, at most northeastern sites in the U.S., the major part of summer FP is sulfates. If CF is assumed to be of local origin because of shorter atmospheric residence times than FP, then there may be a high correlation of FP and CF over time because of changes in weather conditions.[3]

LW97 have overlooked the role of the air pollution and weather data in time series studies. If the long-term seasonal and secular trends are adequately removed from the data set, then the analyses relate *changes* in mortality or hospital admissions rates over a few days to *changes* in probable environmental factors that are known to be related to exacerbation of mortality and illness from cardiovascular and respiratory causes. For the purposes of time series studies, the primary issue is the capability of the monitors to detect *changes* in ambient concentrations to which these populations are exposed. The siting of the monitors is adequate for evaluating changes in $PM_{2.5}$ and in PM_{10} in all 6 cities in SDN, adequate for detecting changes in CF in 5 of the Six Cities, and may even be adequate for characterizing *daily changes* in CF levels in eastern Tennessee.

LW97 also assert that "In joint regressions of two correlated variables with differing exposure measurement errors, the entire effect may be shifted onto the more precisely measured variable (Ref. 6)."

---

[3]Of course, if one were to accept the authors' argument that the ine and coarse fractions of PM have roughly equivalent toxicity, then their arguments regarding exposure suggest it would still be far more effective to implement strategies to reduce fine particles than coarse particles.

EPA agrees that this is well known. The real issues, however, are quantitative: how closely correlated are the variables and their measurement errors, and how large are the errors in real life?

The simulations in Ref. 6 (LW97, 1995a) constitute an extremely unrealistic case. The assumptions on p. 680 are: (1) correlation between predictors (denoted P1 and P2) is 0.96 (almost perfect linear association!); (2) measurement error variance is 0.56 of the actual population variance of P2. The standard deviation is the square root of the variance, so assumption (2) says the standard deviation of the measurement error is about 75% as large as the true variability in P2 -- this is highly implausible in real life. The attenuation expected for a classic least squares regression model if P2 alone were used as a predictor would be to multiply the true regression coefficient on P2 by a factor of :

(true P2 population variance) / (true P2 population variance + measurement error variance)= 1 / (1 + 0.56) = 0.64.

In other words, the assumed measurement error would be expected to deflate the effect size of P2 by about 36%. In the simulation carried out by LW97, the simulated regression coefficients of response on P2 (true) and P2 (measured) were 1.81 and 1.07 respectively, or a ratio of 0.59, suggesting that the simulations were correct. When this is combined with a near-perfect Pearson correlation of 0.96 between P1 and P2, the consequences are also predictable from theory. When the true regression coefficients for P1 and P2 (true) are equal to 1, the estimated coefficients are equal to 1.06 and 0.85 respectively, but not statistically significant because of the extraordinarily high collinearity (r = 0.96) between P1 and P2. When the true regression coefficients for P1 and P2 (with error) are equal to 1, the estimated coefficients are equal to 1.88 and -0.01 respectively, with the P1 coefficient significant but the P2 coefficient not statistically significant because of the extraordinarily high collinearity (r = 0.96) between P1 and P2 and the extremely large measurement error on P2.

EPA concludes that the LW97 simulation is consistent with the mathematical theory fully explained in research monographs such as *Measurement Error Models* by W. Fuller (J. Wiley, N.Y., 1987), and in *Nonlinear Measurement Error Models* by R. Carroll, D. Ruppert, and L. Stefanski (Chapman and Hall, N.Y. and London, 1995). However, EPA staff and expert consultants (e.g., Duncan Thomas) have greatly extended the analyses of the combined effects of predictor variable correlation, measurement error correlation, and measurement error variance (Marcus, 1997). The findings are that substantial bias in coefficient estimates (especially inflation of the estimated size of a real effect, or reversal of its sign) appears to require all three of the following conditions: (1) measurement error size for error-prone predictor variables must be large relative to population variability, at least 50 to 100 percent of the population standard deviation for each; (2) error-prone predictor variables must be highly correlated, e.g., $r < -0.9$ or $r > 0.9$; and, also, (3) measurement errors for these predictors must have a strong negative correlation. The necessary co-occurrence of all three of these extreme conditions is not very likely, and no evidence demonstrating that they have, in fact, co-occurred in any of the PM epidemiology studies relied upon by EPA (including SDN) has been provided to EPA. Thus, while the measurement error issues raised by LW97 (and in their earlier papers) are interesting

and theoretically could be of concern under extremely unusual conditions, EPA finds no compelling basis for rejecting the published PM epidemiology findings relied upon in the Promulgation Decision. See also Appendix D for more discussion of measurement error.

LW97 also indicate that it is reasonable to hypothesize that CF would have been similarly significant in locations other than Steubenville if more reliable CF data had been available there. Again, this is purely hypothetical speculation. The preceding discussion demonstrates a number of ways in which neither the magnitude of the CF measurement error nor the magnitude of the correlation between FP and CF (estimated at 0.69, not 0.96) may be so great as to preclude separating FP and CF effects. It would probably require FP and CF correlations much higher than 0.23 to 0.45 observed in the other five cities to account for the possible misattribution of CP effects to FP as a result of CF measurement error. The simulations of LW97 (Ref. 6) required a 0.96 correlation between predictors (pollutants like FP and CF), and a measurement error of which the standard deviation was 75% of the population variability, to bleed away all of the apparent effect to the perfectly measured pollutant. No such FP vs. CF correlations have been found in the Six Cities, with the possible exception of Steubenville where the measured correlation of 0.69 could indicate that the true correlation is somewhat larger. Differences in measurement error are possible, but it would require measurement errors of enormously larger magnitude between FP and CF to completely attenuate a similar FP effect in the other five Cities. The hypothesis proposed by LW97 does not seem to be quantitatively plausible, although it may be theoretically possible.

LW97 also note that the correlations between FP and CF vary almost in direct proportion to the mean CF value ... This high correlation implies similar behavior in the atmosphere for FP and CF and perhaps common sources (at least in Steubenville). Our hypothesis that FP-CF temporal correlation at a site is controlled by measurement error is consistent with higher CF concentrations leading to higher FP-CF correlation.

Actually, the data from SDN Table 4 show that there is much better correlation of the FP-CF correlation coefficient with mean $SO_4^=$ (r = 0.942), with mean $PM_{2.5}$ (r = 0.914), and with mean PM_{10} (r = 0.899) than with mean CF (r = 0.644). Omitting Topeka, the correlation of the FP-CF correlation is highest for $SO_4^=$ (r= 0.941), but almost identical for the other PM metrics (r = 0.916 for CF, r = 0.912 for PM_{10}, r = 0.905). The FP-CF association is more plausibly related to the common sources of FP and CF in the five Eastern cities, with Topeka CF probably influenced by crustal particles as well as by anthropogenic sources. In Topeka, with the second highest concentration of CF (and by LW97 criteria, second lowest measurement error) in the six cities, the CF-mortality association was negative and non-significant. In summary, the Steubenville FP-CF association proves nothing about the role of CF measurement error.

LW97 also indicate that their Figure 1 implies that (true) ambient FP and CF are <u>not</u> distinct entities (even though some of their sources differ) and that the apparent lack of correlation cited by SDN is simply a result of the increased importance of measurement error at low concentrations. The authors' hypothesis is not confirmed by the data in their Figure 1. As noted above, the only plausible measurement errors identified by LW97 are additive analytical and instrument errors. These measurement errors were not greater than 5 $\mu g/m^3$ in the Six Cities and

were of unknown magnitude in Birmingham and Toronto. Burton et al. (1996; p. 401) estimate uncertainties of 11% for $PM_{2.5}$ and 9.2% for PM_{10}, based on 28 collocated samples. The average absolute measurement error would be about 2 $\mu g/m^3$. The absolute CF error could not be much larger. Even though CF levels were lower than FP levels in Philadelphia in 1992-93, a similar absolute error would amount to a relative error of about 25 to 30%. It is highly unlikely that measurement error of even this magnitude would deflate the CF effect to nonsignificance, if such an effect actually exists. It is more likely that the high FP-CF correlations in Birmingham, Steubenville, and Toronto reflect common anthropogenic sources within each region, such as emissions related to ferrous metal processing operations. The other metropolitan regions with low FP-CF correlations may have a greater diversity of FP and CF sources than these cities.

LW97 broadly generalize in stating that all analyses of the CF data are limited by the inherent CF exposure errors, the corrupting effects of which cannot be removed. Again, actually there are many methods by which measurement error effects can be removed, if necessary. Most of the exposure measurement errors that concern LW97 are of so-called "Berkson" type, for which no further adjustment is needed in classical linear regression models (see also Appendix D). Analytical and instrument errors are classical problems, for which many adjustment methods are available. However, as noted above preliminary assessments of the magnitude of the analytical and instrument errors suggests that they should have had relatively little quantitative effect on the findings of the epidemiology studies being evaluated. In particular, it is highly unlikely that real CF health effects could have been reduced to non-significance, and the CF effects then attributed to FP, by measurement error. As LW97 themselves have shown, this would require at least two highly implausible conditions: (1) measurement error on CF comparable in magnitude to the variability in the population of CF measurements; and (2) an almost perfect true correlation between FP and CF. With correlations among predictors even < -0.9 or > 0.9, a third condition also appears necessary (Marcus, 1997): (3) the measurement errors must also be highly negatively correlated.

LW97 further state that exposure error also obscures the true shape of the underlying exposure-response function. EPA agrees that the distortion of the concentration-response curve is a theoretical possibility. If the response functions were strongly nonlinear, then some distortion may occur for exposure measurement errors of the Berkson type as well as in the classical errors-in-variables model. However, there is little evidence that the response function has more than modest nonlinearity at current levels of PM and other pollutants below the current NAAQS.

E. Review of Morris (1997) Paper on Regression Analysis of Daily Mortality and Air Pollution Using Real and Random Data (Submitted by Roth Assoc. On Behalf of Amer. Auto. Manuf. Assoc., March 1997, IV-D-2243)

Morris (1997) performed a daily time series analysis of mortality data from Manchester, UK, for the years 1988-1992, using BS, SO_2, and NO_2 as air pollution predictors. Weather variables were absolute deviations of temperature from 16 degrees C, and dummy variables for very hot (> 27 degrees) and very cold (< 0 degrees) days. Smoothing approaches were deviations from 19-day moving averages and Fourier series, with ca. 20-day smooths between individual years.

Some regressions were fitted using a "total pollution" variable = $(BS + SO_2 + NO_2)/3$. The author found no statistically significant associations with previous-day pollution, and significant association with 2-day lag temperature deviations. The author also performed a sensitivity analysis by replacing real total mortality residuals from detrended time series with a Poisson noise series, twice, then refitting the model and obtained similarly non-significant results. The author concludes that some weak associations of PM with mortality may be due to chance.

EPA finds that various technical limitations render this study of little value. These include:

(1) probable misspecification of temperature model because a V-shaped function was used, in spite of the author's finding of a *broad* minimum around 16 degrees.;

(2) probable misspecification of extreme temperature model because dummy variables were used;

(3) probable under-specification of weather model because humidity was not evaluated;

(4) probable under-specification of weather model because lag structure were not evaluated, other than a 5- to 7-day lag for the freezing-weather dummy variable;

(5) inadequate evaluation of alternative detrending or smoothing models;

(6) no evident model-building strategy or model specification search, notwithstanding the author's awareness that a variety of lag structures and model specifications have been found optimal in other studies;

(7) many other investigators calculate residuals by removing long-term trends *and* weather effects, and only then assess air pollution effects (e.g., Samet et al., 1996b for Philadelphia), whereas the author appears to have used residuals after removing only long-term trends;

(8) the use of a total pollution variable cannot be justified in view of seasonal differences and possible differences in lag structure and health effects associated with different pollutants;

(9) probable under-specification of pollution models because lag structures were not evaluated; and

(10) only two simulations were carried out, and no effort appears to have been made to evaluate any remaining short-term temporal patterns in the real time series residuals that should have been reproduced in the simulated series.

While the author has presented some novel approaches, there has been so little investigation of plausible alternatives that EPA finds no basis for the author's conclusions.

To amplify further on EPA's concerns, it should be noted that it is not clear why Morris decided to use a couple of random series as simulated mortality data. The reasoning appears to be: (1)

(real) mortality-PM associations were weak and non-significant in this data; (2) random mortality data-PM associations were also weak; and (3) therefore, the observed weak (and non-significant) PM-mortality associations (in this data) are not different from those in PM-random mortality data! This seems to be redundant. The first question is: why take the results from these couple of random series seriously? They could have been non-significantly negative as well. Experiments with random numbers are usually done with many replicates to estimate a range of distribution of certain outcomes. It is pointless to talk about one or two particular trials. These efforts could warrant more serious consideration if 50 researchers had reported weakly positive associations between air pollution and the first two samples, not chosen ones, of randomly generated mortality series.

Consistency or any meaningful structure of weak (non-significant) mortality-PM associations could have been evaluated from this data internally. For example, Morris could have obtained a regression coefficient for today's mortality on tomorrow's air pollution (i.e., lagged the other way). If such coefficients are similar, in magnitude and sign, to that for 1-day lagged pollution, then such associations cannot be taken seriously. It would have been much more constructive if Morris had applied a number of modeling approaches, including ones similar to those used in the past studies, and evaluated the sensitivity of results to different model specification, rather than throwing dice a couple of times.

Unlike many (especially eastern) U.S. cities, the highest PM levels, as measured by BS (a good measure of carbon) in this data occurred in winter, and there is not much variation in BS during summer. The mortality time-series also peak in winter. From the time-series plots, one can also see that the highest influenza peak occurred in the winter when BS was the lowest in the data. Thus, controlling for season and other long-wave components seems particularly important in this data set. The author tried a 19-day moving average and sine/cosine series (what periodicities were used in their sine/cosine series are not clarified), and found this detrending satisfactory. However, it is not clear how sensitive the results could be to specification of the detrending (e.g., different averaging span). Since the PM index here is essentially a winter time pollutant, dividing the data into warm and cold seasons to see if the slope is larger and more significant in winter would be informative. Other modeling specification issues are discussed below.

To circumvent the collinearity of the three pollutants (r= 0.4-0.7), the author also used the average BS, SO_2 and NO_2 [$(BS+SO_2+NO_2)/3$] in the analysis. While the range and unit of these pollutants seem comparable, this is not adequate because the mean of BS is less than half of those for SO_2 and NO_2. The temporal fluctuations of these pollutants are also different: the peaks of BS are more localized in winter, and its distribution is more skewed than other two. Thus, it is unlikely that this operation would extract the common air pollution fluctuations. Because all the air pollution indices showed non-significant mortality associations individually in the regression, multi-pollutant models might not be particularly useful, but sometimes addition of one variable in the regression helps the significance of another if it "clears out" the otherwise unexplained variance. Since the intercorrelation among the pollutants is not overwhelming, multi-pollutant models might have been worth trying.

The basic modeling approach is not clear in this study. A common (and possibly "conservative") approach is to fit the mortality with weather and seasonable variables as much as possible, and then enter the pollution variables of interest. According to the results in Table 2, it appears that the optimal lag of PM (one-day) was determined first. Did this lag show the highest significance for all the pollutants and weather model specification or did the optimum lag for pollution change depending on the model specification? Not enough information is provided to sort out these essential details here.

In addition to extreme temperature indicator variables the absolute deviation of mean daily temperature from 16ºC was used as a temperature variable. Several studies have shown that the lags are different for heat effects (0-1 day) and cold temperature effects (2-4 days). Thus, the deviation variable for heat and cold temperature effects may not be able to adequately model these lagged effects. Recent studies also point out that the slopes are different for heat (steeper) and cold (shallower) effects, although the indicator variables may have taken care of this.

The fact that the BS peaks are mostly localized in winter with a very skewed distribution raises additional modeling concerns, and checking the influence of the extreme values and a winter-only analysis are recommended. Considering the size of the daily death counts (mean-58) and the length of the series (~ 5 years), it is surprising that even the temperature variables are barely significant. This may be indicative of inadequate modeling. More sensitivity analyses of this data by other researchers would be useful.

F. Review of Paper by H.D. Roth & Y. Li on Prague PM Mortality, Submitted by Amer. Auto. Manuf. Assoc. (IV-D-2243)

This report analyzes data from Prague (Czech Republic) for 1986-1994. The PM index used by Roth and Li is described as TSP. SO_2 and NO_x data were available for all years of the study, while CO is only available for 1994. The weather variables used were temperature, humidity, and barometric pressure. Statistical models used were Poisson autoregression and regression on deviations of 19-day moving averages (mortality, and with or without pollution deviations). Lag structures for TSP (and presumably, for other pollutants) were single TSP lag days 0, 1, 2, 3, 4, 5, and averages of days 0+1, 0+1+2,. Roth and Li mention that other studies have used other moving averages, such as lag days 1+2+3 (three preceding days) used by Schwartz for Birmingham (1993). Lags used for weather variables were current day and averages of current day and several preceding days. Roth and Li used dummy variables (deciles) for weather, and quadratic functions.

The authors' findings are that temperature was the strongest predictor of mortality, and that weather lag model affects the magnitude and significance of the TSP lags. Most of the 1920+ models fitted did not show a significant TSP effect. Single-year analyses showed 1/9 positive significant TSP years, 4/9 positive and 4/9 negative non-significant TSP coefficients.

The analysis methods used in this study are very similar to the methods used by Roth and Li in their Birmingham study and, are therefore, subject to the same methodology errors that invalidate the results they obtained in that study. The problems, as discussed earlier for the Roth and Li Birmingham Study, briefly are:

(1) Probable misspecification of Poisson autoregression model.
(2) Use of barometric pressure, almost certainly a covariate confounder that is not used by other investigators for this reason.
(3) Authors do not report use of other TSP moving averages such as 1+2, 1+2+3 etc.
(4) Other 1-day lag weather models not evaluated.
(5) No search for correct model specification.
(6) Quadratic temperature model is probably wrong because the function has the same shape for extremes of hot and extremes of cold.
(7) Dummy temperature model is probably wrong because it assumes level response at hottest 10% of days and coldest 10% of days.
(8) Year-by-year analyses shrink t-statistics to 1/3 of 9-year analysis, everything else being equal.
(9) Not enough information is presented to evaluate pattern of responses, as EPA found using authors' supplemental tables for Birmingham.

Of some particular interest was the indication that TSP effects peaked with the second and third days' levels. Thus Roth and Li should have also evaluated moving averages with lags 2+3, 1+2+3 etc.

Also, Roth and Li noted that the predictive power of TSP was strongest when TSP and NO_x were in the model. This is almost the same finding as Samet et al. (1996b) for Philadelphia. In many studies, inclusion of copollutants may strengthen the estimated PM effect. However, since SO_2 and NO_x are precursors of fine particle component sulfates and nitrates, a model that uses these pathway relationships would be even better.

Overall, what is noteworthy about the paper, as submitted, is what is missing. This includes a discussion about the relative stability of TSP coefficients for current day and for lags 2 and 3 when copollutants are included in the model (Table 4). Also, no information is provided about the results of the weather models. The low values of R-squared in all of the models shown suggests two possibilities: (i) All of the Roth & Li models are inadequate, probably misspecified; and (ii) R-squared is an inappropriate criterion for assessing Poisson regression models. EPA believes that both of these may be occurring.

G. Review of Burnett RT, Dales RE, Brook JR, Raizenne ME, Krewski D. (1997) Association between ambient carbon monoxide levels and hospitalizations for congestive heart failure in the elderly in 10 Canadian cities. Epidemiol. 8:162-167.

The intent of this study, as presented in the authors' introduction, is to attempt to verify the findings of Morris et al. (1995) and Schwartz and Morris (1995) regarding an association

between exposure to ambient carbon monoxide and admissions to the hospital for congestive heart failure. Morris and colleagues examined the association between daily changes in levels of several gaseous air pollutants and daily hospital admissions of the elderly for congestive heart failure in seven U.S. cities; only CO showed a consistent association. Schwartz and Morris conducted a similar study in Detroit, but the analysis was expanded to include other categories of cardiovascular disease. For congestive heart disease, significant associations were found for both CO and PM_{10} in single pollutant models, as well as when combined in a two-pollutant model. For hospital admissions of all cardiovascular diseases combined, the authors report that PM_{10} showed the most consistent associations. Significant increases in hospital admissions for ischemic heart disease were found with PM_{10}, CO and SO_2 in single pollutant models, but only the association with PM_{10} remained signicant in two-pollutant models.

Burnett et al. (1997) obtained admissions data from hospitals in 10 large Canadian cities for approximately an 11-year period, and selected for analysis the hospital admissions for congestive heart failure among patients over the age of 65 years. A random effects relative risk regression model was used to evaluate relationships between admissions and pollutant levels; the models included a 19-day linear filter to remove long-term and seasonal trends and control variables for day-of-week and weather factors. Separate analyses were conducted for each month of the year, to control for seasonal variability between cities, and pooled estimates of the log-relative risk for the ten cities for each month were obtained. The common log-relative risks reported in the paper are weighted averages of the monthly log-relative risks.

EPA believes that even combining month-by-month analyses across different cities may tend to reduce the power of the analysis to detect effects, since differences in the effects of pollution and weather may depend on city-specific characteristics as well. It is not clear that the use of the site as a random effect completely compensates for possible site-specific differences in the shape of the dose-effect relationship.

CO, NO_2, SO_2, ozone, and COH were used in single-pollutant models, and significant associations were found for same-day CO, NO2 and COH; NO2 lagged one day was also significantly associated with admissions for congestive heart disease. The correlation coefficients for these three pollutants ranged from 0.50 to 0.56. The three pollutants were insensitive to adjustments for weather or seasonal effects, but COH lost statistical significance when the four other co-pollutants were included in the model. The authors conclude that, of the five pollutants, CO is the strongest predictor of hospitalization rates for congestive heart disease, and that this finding is consistent with the results of Morris et al. (1995) and Schwartz and Morris (1995).

While the analysis of data on a monthly basis could result in a loss of statistical power to detect associations in the sub-analyses, the results of this study are consistent with the findings of previous studies regarding associations between congestive heart failure and exposure to CO. The association found here was with unlagged CO concentration, which is coherent with the known biological action of CO in the formation of carboxyhemoglobin and the resulting additional stress on the heart.

These results are also not inconsistent with the many previous studies that have shown associations between PM exposure and death or hospital admissions for cardiovascular disease. Burnett et al found a positive association between unlagged COH and admissions for congestive heart disease that disappeared with the addition of CO, NO_2, SO_2 and ozone to the model. In the previous study, Schwartz and Morris (1995) found associations with PM_{10} and CO that were independent of one another; the correlation coefficient for the two pollutants in this study was 0.30. While little data are available for associations between pollutants and congestive heart failure, it appears that it may be difficult to separate the effects of CO and PM for this health effect. PM may also be associated with transportation-generated sources. As stated by Burnett et al: "Ambient carbon monoxide concentrations are highly correlated with nitrogen dioxide and the coefficient of haze. All three pollutants are generated primarily from common transportation sources. In univariate models, all three pollutants displayed positive associations with congestive heart failure. Thus, carbon monoxide may be acting as a marker for pollution from transportation sources in general."

PM has been found to be strongly associated with admissions or death from cardiovascular disease in many studies, as summarized in the CD and Staff Paper. As discussed in the responses on toxicological studies, a number of new studies that were not published prior to the closure of the CD have indicated the presence of a relationship between exposure to ambient PM and cardiovascular disease in animals. Under the hypothesis proposed by Seaton et al. (1995), acidic ultra-fine particles cause inflammatory changes that result in changes in blood coagulability that may increase susceptibility to acute episodes of cardiovascular disease. One recent study indicates that blood viscosity was increased with pollution exposures during a pollution episode in Germany (Peters et al., 1997), and some new toxicological studies in animals have found evidence for rapid failure of the heart due to conduction-related arrhythmia (Watkinson et al., 1997).

In summary, EPA does not agree that the results of Burnett et al. (1997) disprove an association between PM and cardiovascular health effects. This study evaluated the relationship between congestive heart disease hospitalization and ambient CO, and found a positive association. Co-pollutants were also used in the analysis, and some evidence of association was found with NO_2 and COH, but CO was found to have the strongest association. These results do not in any way conflict with the findings of increased death or hospitalization for cardiovascular disease in numerous studies.

II. SUMMARY OF APHEA PROJECT RESULTS

The APHEA (Air Pollution and Health: a European Approach) project, supported by the European Union Environment 1991-1994 Programme, was conducted to investigate short-term health effects of exposure to low or moderate levels of ambient air pollutants. The APHEA project studies were carried out by 11 research teams in 15 cities (Amsterdam, Athens, Barcelona, Bratislava, Cracow, Helsinki, Koln, Lodz, London, Lyon, Milan, Paris, Poznan, Rotterdam and Wroclaw) in which air pollutant data had been collected for at least five years. A

series of APHEA study reports was published in 1996, after the closure of the EPA PM Criteria Document (PM CD), and thus the findings were not reviewed and discussed in the PM CD.

Nevertheless, a number of public commenters drew attention to the "negative" APHEA studies as evidence against consistency of PM-associated health effects. In response, EPA believes that it is important to summarize the findings of all of the APHEA studies, in addition to the findings of two meta-analyses that have been recently published. These studies, and this summary of the studies' findings, have not been reviewed by the Clean Air Scientific Advisory Committee, and thus they are not included in the scientific support for the EPA's rulemaking decision. Probably the most important fact to emphasize here is that, contrary to the selective allusion by some commenters to "negative" APHEA results, in fact (as discussed below), the results of these studies and subsequent meta-analyses generally support EPA's finding of consistency with regard to PM epidemiologic results derived from different research efforts in the U.S. and internationally. Key features of the various APHEA studies for individual European cities are summarized in Table C-3.

In this discussion of APHEA study results, it is essential to note that different indicators of PM concentration were used in the various European countries, and that none of the studies explicitly involved ambient measurements of PM_{10}, $PM_{2.5}$ or $PM_{10-2.5}$ (the size-specific PM indicators of most direct relevance to the present EPA PM NAAQS Promulgation Decision). Most of the studies used measures of TSP (Bratislava, Helsinki, Koln, and Milan) or BS (London, Paris, The Netherlands, Barcelona, Athens, Cracow, Lodz, Poznan, and Wroclaw). In addition, monitoring systems in France measured PM_{13} (Dab et al., 1996; Zmirou et al., 1996), while a measure of PM_7 was used in the German study (Spix and Wichmann, 1996). The specific methods used to measure BS, TSP or other forms of PM were also quite different from city to city (See Table C-4). The variation in PM monitoring methods and PM indicator use in the APHEA studies is an important caveat to the interpretation of the findings of these studies, whether the findings be "negative" or "positive" for associations between PM exposures and health. Full evaluation of the studies and their ultimate implications will not be possible until additional information (beyond that reported in the published APHEA papers) is obtained on the monitoring methods used, including information on calibration and quality assurance procedures.

In addition to the varied PM measurement methods used, all APHEA studies included ambient sulfur dioxide measurements. Concentrations of NO_2, O_3 and CO were also available for some, but not all APHEA studies. One important feature of the APHEA project was establishment of a standardized statistical analysis protocol, so that the findings from the different cities could more readily be compared and used in meta-analyses. The analytical methods used included adjustment for long-term, seasonal, or day-of-week trends, meteorological factors, the presence of influenza epidemics or other such confounders, and autocorrelation in the data sets. The APHEA studies used daily counts of mortality or hospitalization as health outcomes, and meta-analyses have been published for the results of studies on mortality for all causes, excluding accidents (Katsouyanni et al., 1997), and for hospitalization for chronic obstructive pulmonary disease (COPD) (Anderson et al., 1997).

Table C-3. Summary of APHEA Individual City Study Findings

Study	Independent Variable(s)	Dependent Variable(s)	Findings and Conclusions
Bratislava, Slovak Republic, 1987-1991, Bacharova, et al.	Particles: TSP Other: SO_2, Mean TSP ranged from 77.84 to 101.79 ug/m^3 by season (highest in winter)	Daily deaths: - total (approximately 10/day), - total minus deaths from external causes, - respiratory, - respiratory infections, - obstructive lung diseases, - cardiovascular, - ischemic heart disease, - lung cancer. Control = digestive causes	* No significant effects reported for either SO_2 or TSP * However, Table 5 lists a significant RR for SO_2 with digestive diseases mortality (RR 1.035, CI 1.01-1.05), and also an RR of 1.008 for TSP with total mortality with an impossible CI of 0.96-0.99.
Athens, Greece, 1987-1991, Touloumi, et al.	Particles: BS Other: SO_2, CO Mean BS 84.4 ug/m^3, range 9-333	Total death count, average 37.2/day	* RR for 100 ug/m^3 increase in BS(1 day lag) of 1.05 (CI 1.03-1.08) * Significant associations also found for SO_2 and CO * Significant interaction found for SO2 with season ($p<0.05$) but not BS or CO
London, England, 1987-1998, Anderson, et al.	Particles: BS Other: ozone, NO_2, SO_2 mean (range) BS 14.6 (3-95) ug/m^3	Daily mortality counts: all cause excluding accidents, respiratory cardiovascular	* For BS, (24-hr, lag 1), predict increase of 2.5% all-cause mortality in the warm season, 1.6% increase in the cool season, and 1.70% all year; all associations statistically significant. * Associations of BS with deaths from respiratory and cardiovascular diseases not significant. * Significant associations also found for ozone for all seasons, and for SO_2 and NO_2 during the warm season * BS and ozone remain significant in two-pollutant models for all-cause mortality; the effects of both are relatively unchanged in the warm season, while BS effect was doubled, but remained nonsignificant, in the cool season.

Paris, France, 1987-1992, Dab, et al.	Particles: BS, PM_{13} Other: SO_2, NO_2, and ozone Mean (5th and 95th percentiles) in winter BS 39.9 ug/m^3 (12.6-137.6); PM_{13} 54.4 ug/m^3 (25.0-151.0); Summer levels lower	Mortality: only respiratory codes used (mean 9/day). Hospitalization: respiratory disorders (mean 79/day).	* Respiratory mortality was significantly associated with PM_{13} (0-1 day) (RR 1.168, CI 1.041-1.310); positively associated with BS but not significant (RR 1.071, CI 0.98, 1.18) * Admissions to hospital for all respiratory causes was significantly associated with BS (0-day) (RR 1.041, CI 1.007-1.075), PM_{13} (0-day) (RR 1.045, CI 1.004-1.087); no associations with admissions for COPD or asthma * Significant associations also found for respiratory mortality with SO_2, hospitalization for COPD and asthma and SO_2, asthma hospitalization for NO2, and there was a negative association between asthma hospitalization and ozone
London, England, 1987-88 and 1991-92, Ponce de Leon, et al.	Particles: BS Other: ozone, NO_2 and SO_2 Mean (5th and 95th percentiles): BS 14.6 ug/m^3 (6-27)	Admissions to hospital for respiratory causes (mean 125.7 for all ages)	* Only sporadic associations were found with NO_2, SO_2 and BS; most consistent associations found with ozone, * Only significant finding for BS was a negative association for 0-14 yo, all-year (RR 0.9815-0.9993)
Helsinki, Finland, 1987-1989, Ponka and Virtanen	Particles: TSP Other: SO_2, NO_2, O_3 mean TSP 76 ug/m^3	hospital admissions for asthma via emergency wards in Helsinki (2.21 admissions per day) Admissions for diseases of digestive system used as control.	* Significant findings only for ozone and SO_2
Amsterdam and Rotterdam, The Netherlands, 1977-1989, Schouten, et al.	Particles: BS (only for 1986-9). Other: SO_2, NO_2, O_3 24-hour BS mean (95th percentile) in winter: 14 (50) ug/m^3 in Amsterdam and 28 (72) ug/m^3 in Rotterdam (summer levels similar)	Emergency hospital admission data for respiratory diseases (respiratory, COPD and asthma), using 15-64 and 65+ age groups. Means for Amsterdam and Rotterdam (respiratory, all ages) were 6.70 and 4.79 per day	* Black smoke findings were mixed. In Rotterdam, daily mean BS (1-day lag) was associated with respiratory admissions in 15-64 y.o. group (RR=1.374, CI=1.091-1.730), but other findings were sometimes positive, sometimes negative.

Koln, Germany, 1975-85, Spix and Wichmann	Particles: TSP, PM_7 from neighboring city Other: SO_2 NO_2 TSP median of 68 ug/m^3 all-year, and 69 (max 304) for Dec-Feb	total mortality (median 29 per day)	* PM_7 predicted a 2% increase in mortality between the 5th and 95th percentiles (lag 1); the association had borderline significance ($0.05<p<0.10$). The RR for TSP (lag 1) was also 1.02, but not statistically significant. * SO_2 also associated with mortality; no associations with NO_2
Barcelona, Spain, 1985-1991, Sunyer, et al.	Particles: BS Other: SO_2, ozone, NO_2 Winter median BS 49.7 ug/m^3 (11.4-66.7), Summer median BS 35 ug/m^3 (10.6-125.6)	Mortality: -total, -total for 70+ years of age, -cardiovascular, and -respiratory causes Medians of 48 and 43 deaths/day in winter and summer; with only 5 and 4 deaths/day for respiratory in winter and summer	* BS was positively associated with all mortality types, though the associations were not significant with respiratory causes, and only the associations during winter were statistically significant for the other types. For a 100 ug/m3 change in BS for the full-year data (1-day lag for all but 3-day for respiratory): total RR 1.070 (1.029-1.112), total > 70 yr RR 1.063 (1.016-1.113), cardiovascular RR 1.093 (1.036-1.153), respiratory RR 1.097 (0.990-1.215) * Significant associations also found for SO_2, NO_2 and ozone
Milan, Italy, 1980-1989, Vigotti, et al.	Particles: TSP Other: SO_2 mean TSP 139.0 (3.5-529.5) ug/m^3	- Daily deaths from respiratory causes (mean 2.9/day) - Daily hospital admissions for respiratory causes (mean 11.3/day for 15-64 yo and 8.8/day for 65+ yo)	* unlagged TSP associated with respiratory mortality (RR 1.12, CI 1.02-1.23), with a seasonal interaction term ($p<0.10$) indicating a trend toward stronger effect in the warm season. A significant seasonal interaction term also with cumulative exposure, with a stronger effect in the warm season (RR 1.39, CI 1.15-1.67) * 2-day lagged TSP was associated with respiratory hospitalizations (RR 1.05 CI 1.00-1.10) in the 15-64 yr age group, and the effect was greater during the warm season (RR 1.08, CI 1.00-1.17); positive associations also in the 65+ yr age group that were significant in the cool season. * Significant associations also found for SO_2

Lyon, France, 1985-1990, Zmirou, et al.	Particles: PM_{13} Other: SO_2, NO_2 mean PM_{13} 38.05 (2.67-179.81) ug/m^3	Daily counts of mortality: - total minus external causes deaths (6.43 per day mean), - respiratory (0.40/day), - cardiovascular (2.04/day) - digestive diseases (0.38/day) deaths used as control.	* Positive associations with PM_{13}, but only significantly associated with respiratory deaths (0-day, RR 1.04, CI 1.00-1.09); RR's for cardiovascular and respiratory mortality were greater during the cold season. * Significant associations found with SO_2 * No significant associations with NO_2 or O_3
Cracow, Lodz, Poznan and Wroclaw, Poland 1977-90 (full study), Wotyniak and Piekarski	Particles: BS Other: SO_2 Medians, (10th and 90th percentile) for BS (ug/m^3): - Cracow 73.3, (25.6, 246.7), - Lodz 57.3, (20.3, 150.5), - Poznan 34.0, (9.3, 92.0), - Wroclaw 54.3, (25.7, 141.3)	Mortality: - total deaths minus external causes, - respiratory diseases, - circulatory, - digestive system (control)	* Mixed results for both pollutants, including a positive association between BS and digestive diseases in Cracow * Significant positive associations found for BS with total mortality in both Cracow and Lodz, but significant negative association found with respiratory diseases in Wroclaw

APHEA Methodology, as summarized from Katsouyanni et al (1996):

(1) Pollutants were analyzed in single-pollutant models, and interactions or confounding were tested by stratified analysis in high/low pollutant categories (100 ug/m^3 cut-off for SO2 or PM), and also adding an interaction term for the high/low pollutant variable and the pollutant being tested

(2) Model-building done first using linear regression of log-transformed health variables (adding some value if there are counts of 0 in the data). Final models then tested in Poisson models with autoregression (autoregressive terms up to 4) and allowance for overdispersion -- RR's came from these final models.

(3) Time series plots were used to identify trends or seasonality, and residuals from the predicted time series were plotted to assess appropriateness of correction for trends, seasonality and epidemics

(4) Model building steps:

a) seasonality addressed with sinusoidal (sine/cosine) functions for cyclic patters

b) long term trend, if there was a systematic change in health outcome, was addressed with the addition of a variable, and the square of the variable if needed, for the number of days in the study, and if there were year-to-year fluctuations, dummy variables were added for each year

c) after addressing seasonal trends, residual-residual plots for mortality and temperature or humidity were used to select appropriate meteorologic terms, including: (1) linear), (2) linear and quadratic terms, (3) double quadratic, or U-shaped curve around some turning point, and (4) third-order polynomial or other. Various lags were tested, and the same lag was used for temperature and humidity in a single data set

d) dummy variables were added for day of week and holidays (if needed)

e) indicator variables were added for occurrence of influenza epidemics

f) air pollutants were added after inspection of time plots for residuals -- lags of 0 to 3 were tested for all pollutants except ozone, for which up to 5- day lags were tested. Cumulative lags of 0-1, 0-2 and so forth were also tested

Table C-4. Summary of PM Measurement Methods Employed in APHEA Studies

Study Senior Author	PM Index	PM Measurement Method Used in Given City
Anderson	BS	British standard method, 24-hr average of 4 sites in London
Bacharaova	TSP	Filter gravimetic, 24-hr average of 9 stations in Bratislava
Ballester	BS	Reflectometric method, 24-hr average of 3 stations in Valencia
Dab	PM_{13} BS	Reflectometric method for BS, radiometry for PM13, 24-hr avg over 15 stations in Paris
Ponce deLeon	BS	British standard method, 24-hr avg. of 8 stations in London
Ponka	TSP	Method not described, 4 stations in Helsinki, 4 in suburbs.
Schouten	BS	OECD method, 24-hr average, stations in Amsterdam, Rotterdam
Spix	TSP $PM_{7.1}$	Gravimetric, 24-hr. Avg. 3 times/week, 5 sta. in Koln (Cologne). $PM_{7.1}$ from Dusseldorf, 40 km north.
Sunyer	BS	OECD method, 24-hr avg of 7 stations in Barcelona
Toulomi	BS	OECD method, 24-hr avg of 3 stations in Athens region
Vigotti	TSP	24-hr avg., beta attenuation, 2 sites for TSP in Milan
Wojtyniak	BS	Reflectometric method, 24-hr avg. 3 sites each in Wroclaw, Cracow, Lodz, Poznan
Zmirou	PM13	24-hr avg., beta-ray AAS (similar to PM_{10}), 3 sites for PM (one outlier station deleted) in Lyon

As correctly noted by some commenters, not all of the APHEA studies found statistically significant positive associations between PM and health outcomes (mortality, hospital admissions). The study in Bratislava, for example, found no effects for either TSP or SO_2, with the only significant finding being the unlikely positive association between SO_2 and death from digestive causes (the control measure) (Bacharova et al., 1996). In the London study using hospitalization data, the only statistically significant result for BS exposure was a negative association for respiratory causes in the 0- to 14-year-old age group (Ponce de Leon, et al., 1996). Also, the associations found between mortality and both SO_2 and BS in the four Polish cities were quite varied; and there were positive associations between BS and total mortality in Cracow and Lodz, while there was a negative association with respiratory deaths in Wroclaw (Wojtyniak and Piekarski, 1996). There were also no significant associations between TSP and hospital admissions for asthma in Helsinki (Ponka and Virtanen, 1996), and mixed results for BS exposure with hospital admissions for respiratory causes in the two cities from The Netherlands (Schouten et al., 1996).

However, in a number of the APHEA studies, especially those conducted in Western European cities, there were associations between PM measures and health effects that are qualitatively consistent with those found in North American studies. BS exposure was significantly associated with all-cause mortality in London, and there were positive but non significant associations with cardiovascular and respiratory deaths (Anderson et al., 1996). Positive associations were also found between BS and total, cardiovascular or respiratory mortality in Barcelona, though the associations with respiratory mortality and the associations found during the summer months were not statistically significant (Sunyer et al., 1996). In Paris, PM_{13} was more strongly associated with mortality or hospitalization for all respiratory causes than the other pollutants considered, and the estimated relative risk for this association was 1.17 (95% CI 1.04, 1.31) (Dab et al., 1996). Similarly, in Lyon, PM_{13} was also positively associated with total, cardiovascular and respiratory mortality with only the association with respiratory mortality being of borderline significance (relative risk 1.04, 95% CI 1.00, 1.09) (Zmirou et al., 1996). A 2% increase in total mortality was estimated for a 5th- to 95th-percentile change in PM_7 concentration in the German study (Spix and Wichmann, 1996) and a 5% increase in total death count was estimated for a 100 $\mu g/m^3$ increase in BS in Athens (Touloumi et al., 1996). Vigotti et al. (1996) estimated a 5% increase in hospitalization for respiratory causes, and a 12% increase in respiratory mortality, for a 100 $\mu g/m^3$ increase in TSP in Milan.

Table C-5. Summary of Meta Analyses of APHEA Studies in Multiple European Cities

Mortality meta-analysis, Katsouyanni, et al. (1997)	Particles: BS in 8 cities, estimated PM_{10} in 6 cities (based on PM_7, PM_{13} and TSP converted-to-PM_{10}) Other: SO2 in all 12 cities	All-cause mortality	* Overall effects for 50 μg/m^3 increase in pollutants (1-day lag): RR for BS 1.013 (95% CI 1.009, 1.017) and RR for PM10 1.022 (95% CI 1.013, 1.031); all associations statistically significant. * 50 μg/m^3 increase in PM indicators associated with increased mortality in Western European cities: a 3% increase for SO_2 , a 3% increase for BS, and a 2% increase for PM_{10} (1-day and 5-day average lags similar) * 50 μg/m^3 increase in PM indicators also associated with increased mortality in Eastern European cities: a 1% increase for SO_2 and BS; also a 4% increase in mortality for PM_{10} in Bratislava * Similar results found for each pollutant when subset analysis was done for upper and lower ranges of the other pollutant, indicating independence of BS and SO_2
COPD hospitalization meta-analysis Anderson, et al. (1997)	SO_2, BS, TSP, NO_2 and ozone available for 5 cities; only SO_2 and TSP in Milan	COPD hospital admissions	* Summary RR's for 50 μg/m^3 increases in single-day lags (0- or 1-day, depending on the study): BS 1.04 (95% CI1.01, 1.06); TSP 1.02 (95% CI 1.00, 1.05)

As noted earlier, the APHEA project was designed to facilitate the use of meta-analytical methods on the results from the different research teams. Two such analyses have been published recently (as summarized in Table C-5). One meta-analysis of the APHEA studies found significant increases in total mortality with increases in BS, SO_2, and estimated PM_{10} (Katsouyanni et al., 1997). Daily counts of total mortality were obtained in twelve of the cities (Athens, Barcelona, Bratislava, Cracow, Koln, Lodz, London, Lyon, Milan, Paris, Poznan, Wroclaw), and BS measurements were made in 8 of these cities. Although PM_{10} was not directly measured in any of the cities, it was derived from TSP data using the formula [PM_{10}=TSP x 0.55], and the measures of PM_7 and PM_{13} were assumed to approximate PM_{10}. Such estimated PM_{10} levels were thus used for 6 cities included in the PM_{10} meta-analysis. The effect estimates from the cities were weighted based on the variance of the estimates, and the resulting estimated pooled relative risks are; for a 50 $\mu g/m^3$ increase in each pollutant (lagged one-day) are: 1.013 (95% CI 1.009, 1.017) for BS and 1.022 (95% CI 1.013, 1.031) for PM_{10}. The second study, a meta-analysis of studies using COPD hospitalization data from 6 cities (Amsterdam, Barcelona, London, Milan, Paris, Rotterdam), also found positive associations with BS and TSP (Anderson, et al., 1997). For increases of 50 $\mu g/m^3$ in BS or TSP (lag periods up to 2 days), respectively, relative risks for COPD hospital admissions are: 1.035 (95% CI 1.003, 1.027) and 1.022 (95% CI 0.998, 1.047). An additional finding of note in the meta-analyses is that the effects of SO_2 and PM are apparently independent of one another.

On balance, the overall initial reports from the APHEA studies, and meta-analyses of these studies support EPA's position that there is consistency in the findings of PM-related health effects from studies conducted in different settings and using different methods. The relative risks found for mortality or hospitalization in the meta-analyses of the European studies are not very different from those found for studies conducted in the U.S. or Canada. It is, again, important to emphasize that the findings are based on measurements of PM using several different indicators and different PM monitoring methods in the APHEA study cities. The results from the initial APHEA studies on health effects of based on optical PM measures must be viewed with caution, and the same is true for meta-analyses using these findings. However, it can also be noted that, in the cities where PM_7 or PM_{13} measurements were available, positive associations were found with mortality from all causes (excluding accidents), cardiovascular disease, and respiratory disease, and with significant associations for respiratory mortality in France (Dab et al., 1996; Zmirou et al., 1996) and with all-cause mortality in Germany (Spix and Wichmann, 1996). In these studies, in which the PM measurements were made with a gravimetric method analogous to the methods used to determine PM_{10} concentrations, the relative risks ranged from a 2% increase in total mortality with a 5th to 95th percentile change in PM_7 (Spix and Wichmann, 1996), to a 17% increase in respiratory mortality with a 100 $\mu g/m^3$ change in PM_{13} (Dab et al., 1996).

In this initial series of publications, several APHEA researchers noted two additional concerns that may have influenced the results of their studies. Research teams from Poland (Wojtyniak and Piekarski, 1996) and Helsinki (Ponka and Virtanen, 1996) suggested that the core analytical model design of the APHEA project may not have dealt well with certain patterns in the data sets from these two studies. Ponka and Virtanen discuss the analytical problems in some detail in their discussion, and noted that unexplained cyclic patterns remained in the data even after numerous attempts to address seasonal and other trends with the APHEA core modeling

protocol; it is possible that there are climatic or seasonal differences between cities that were not well addressed in the APHEA analytical procedures. In addition, there may be differences in socioeconomic conditions among the cities. As hypothesized by Wojtyniak and Piekarski (1996): "under different socioeconomic conditions and mortality levels increased because of other risk factors related to lifestyle, diet, occupational exposure, indoor environment etc, the outdoor air pollution level plays a smaller role as a health hazard for the general population." While not a shortcoming of the project, several APHEA project researchers also observed that the chemical nature and size distribution of particulate pollution in the different European cities may differ from those in U.S. or Canadian cities. The differences in aerometry and the substantial differences in location and strength of primary PM emissions sources in central and eastern Europe as compared to western Europe or the U.S. might well explain the different results in these unique areas. Consequently, integration of these results would involve comprehensive examination of the various PM instruments used, monitor siting in relation to sources, mass calibration procedures and other aspects of these studies

In summary, a number of commentors have argued that the scientific literature on health effects of PM exposure is not consistent, but is replete with confusing and contradictory results, and selected APHEA studies have been cited to support this claim. After reviewing the complete set of initial reports from the APHEA project, EPA does not believe this to be the case. While the results of the APHEA analyses of PM-related health effects must be interpreted with caution, it appears that the studies are generally supportive of a small but significant association of ambient PM exposure with increases in mortality and hospitalization.

III. SELECTED OTHER RECENT STUDIES OF PM-RELATED HEALTH EFFECTS PUBLISHED FOLLOWING CLOSURE OF THE CRITERIA DOCUMENT

As discussed in sections I and II, numerous commenters submitted copies of recently available studies not assessed in the PM CD (for example, Davis et al., 1996; Moolgavkar et al., 1997) that fail to find associations between PM exposure and health effects; EPA analyses and discussion of such studies are provided in preceding sections. On the other hand, as discussed in the Preamble to the Final Rule (Section II.B.2.b.), a number of other studies have been published since the closure of the CD and Staff Paper that find positive, statistically significant associations of adverse health effects with exposure to PM. EPA has conducted a brief provisional review of the scientific literature and obtained copies of a number of new studies on the health effects associated with PM_{10} or fine particles. This review cannot be considered a complete literature review, and, while the papers have been individually peer-reviewed, this summary has not been subjected to peer review or the public comment process, and it has not been scrutinized by the CASAC, as is done in the development of the CD and the Staff Paper. However, it is striking that positive, statistically significant results are presented in all of the studies that used fine particles as the PM indicator, and in most of the studies using PM_{10} as a PM measure.

In the brief literature review, five new fine particle (FP) studies were found. Of these studies, Delfino et al. (1997) and Romieu et al. (1996) used $PM_{2.5}$ measures, Linn et al. (1996) used PM_5, Neas et al. (1996) used $PM_{2.1}$, and Peters et al. (1997c) used $PM_{2.5}$ and three smaller-than-$PM_{2.5}$ categories (all also generally obtained data for larger particle measures, such as PM_{10}.) One

study (Romieu et al., 1996) was conducted in Mexico City, while Peters et al. (1997c) was conducted in Erfurt, East Germany, and the concentrations found in those areas were fairly high. However, the other studies were conducted in Canada or the U.S., and the levels of air pollutants in those studies were generally well below current U.S. air quality standards. In all five FP studies, significant associations were found between fine or ultrafine particle concentrations and adverse health effects. These studies are briefly outlined below:

- Delfino et al. (1997) studied associations between emergency room admissions for respiratory causes and air pollutants. The authors found ozone to have the strongest association with ER visits for data collected in 1993. There were also significant associations with PM, and the strength of the associations with the different PM measures were ranked as follows: $PM_{2.5}$>PM_{10}>sulfates.

- Linn et al. (1996) used concentrations of pollutants from ambient monitors, local (in-school) and personal monitors and, as the health measure, repeated pulmonary function tests and symptom reporting from schoolchildren. The authors conclude that PM_5 was more strongly associated with lung function in schoolchildren (more so than co-pollutants O_3 or NO_2).

- Neas et al. (1996) conducted a panel study in State College, PA, obtaining daily reports of lung function (PEFR) and symptoms from schoolchildren, along with concentrations of several pollutants and daily counts of pollens and mold spores. Decreased lung function was associated higher concentrations of all of the particle measures (though none achieved statistical significance), but there were positive, statistically significant associations between reported symptoms and $PM_{2.1}$.

- Peters et al. (1997c) conducted a panel study, obtaining daily reports of PEFR and symptoms from adult asthmatics, while obtaining ambient concentrations of fine and ultrafine particles. Significant associations were found with most fine particle subsets and PEFR or symptoms, but for PEFR the stronger associations were found with the ultrafine particle measures.

- Romieu et al. (1996) conducted a panel study using children in Mexico City, and obtained concentrations of most criteria pollutants. It was found that PEFR was associated with $PM_{2.5}$, PM_{10} and ozone, but the $PM_{2.5}$ associations were the least sensitive to the inclusion of other pollutants in the models. Significant, positive associations were also found between $PM_{2.5}$ concentrations and the presence of cough, phlegm, and lower respiratory illness, while the association with wheezing was positive but not statistically significant. Again, based on changes found in two-pollutant models, $PM_{2.5}$ appeared to be less sensitive to the inclusion of other pollutants in the model than was ozone.

In addition, a number of studies have been published since closure of the CD that studied health effects associated with PM_{10}. Seven recent studies report positive, statistically significant associations between PM_{10} and adverse health effects. The health measures used in these studies

include ER admissions for asthma (Lipsett et al., 1997), mortality or admissions to hospital for numerous causes (Woodruff et al., 1997; Wordley et al., 1997), outpatient respiratory visits (Choudbury et al., 1997), and daily changes in lung function or symptoms (Scarlett et al., 1996; Peters et al., 1996, 1997a, and 1997b). Several individual observations drawn from the studies' conclusions are:

- Lipsett et al. (1997) found significant associations of ER admissions for asthma with PM_{10} during periods of colder temperatures -- in the 20, 30 and 41-degree categories -- the major source in the area (Santa Clara County, California) is wood smoke;

- Woodruff et al. (1997) studied neonatal mortality and associations with PM exposure during the first two months of life, using data from 86 U.S. metropolitan statistical areas. Positive associations were found with total mortality and deaths from respiratory causes and sudden infant death syndrome.

- Wordley et al. (1997) conducted subset analyses by successively removing PM_{10} data above 70, 60 and 50 $\mu g/m^3$, and found no evidence for a threshold in the associations with hospital admissions (all respiratory, bronchitis, asthma, pneumonia, and cerebrovascular) and deaths (all causes, circulatory, and COPD) in Birmingham, UK;

- Scarlett et al. (1996) found significant changes in FVC and $FEV_{0.75}$ across the PM_{10} concentration range in Southeast England, and these effects were unchanged by addition of ozone or nitrogen oxides to the models;

- Peters (1997a) found associations with sulfates, PM_{10}, SO_2 and particle-strong acidity, but sulfates were found to be the strongest predictor of respiratory symptoms in children; this study was conducted in the Czech Republic, in an area with high levels of air pollutants, especially SO_2.

One additional recent publication has reported an association between plasma viscosity in human volunteers with the occurrence of an air pollution episode in Germany in 1985 (Peters et al., 1997d). During the air pollution episode with unusually elevated concentrations of both TSP and sulfur dioxide (maximum levels of 176 and 238 $\mu g/m^3$, respectively), clear increases were found in viscosity. When analyses were done for individual pollutants, the associations of plasma viscosity with TSP were positive, but not statistically significant over the limited duration of this study. In discussing their results, the authors of this study state that "the analyses presented here support plasma viscosity as an intermediate step for both particulate air pollution and carbon monoxide, which have been found to be associated with hospital admissions for cardiovascular diseases."

In closing, it must again be emphasized that this provisional summary represents only a brief overview of some findings reported in the recent literature, and it is not intended to serve as a supplement to the CD. This summary of new studies has not undergone the detailed and extensive review process entailed in the development of a CD, and it has not been discussed by

the Clean Air Scientific Advisory Committee. With those qualifications in mind, this summary is provided in response to some commenters's arguments that findings from new studies are contradictory to those conclusions of the CD and Staff Paper with respect to the consistency and coherence of the available epidemiological information on PM. All of the new studies summarized here using fine particle measures, and most studies using PM_{10} as the PM indicator, report significant associations between PM exposure and adverse health effects.

IV. SUMMARY OF NEWLY AVAILABLE HUMAN, ANIMAL, AND *IN VITRO* TOXICOLOGY STUDIES RELATED TO PM

In addition to the epidmiological studies discussed in the previous two sections of this appendix, a number of recent toxicologic papers have been published, accepted or appear in proceedings that involve exposure to concentrated ambient fine particles, PM emissions, or PM constituents that may be relevant to evaluating the potential mechanisms of toxicity for PM. Some of these studies have begun using animal models that are intended to reflect the kinds of human populations observed to be sensitive to PM in the epidemiologic studies. Section 11.10 and Section 13.5.2 of the 1996 PM CD discuss a number of potential pathophysiological mechanisms by which PM matter deposited in the lung could have adverse health consequences. These included pulmonary inflammatory responses, exacerbation of lung disease, cardiac effects such as arrhythmias, vascular effects such as altered clotting mechanisms, and damage to lung cells. In addition to acute injury, release of various mediators involved in the cellular repair processes can lead to chronic changes in the lung. Although not available at the time the PM CD was completed and therefore not considered in the final Promulgation Decision, there are numerous recently-available toxicological studies that add to the understanding of the pathophysiological mechanisms of PM effects; this section provides a brief summary of a number of the recent studies of interest.

Several investigators have recently reported results of studies of exposure of rats to PM mixes from combustion sources, in this case residual oil fly ash (ROFA) and fuel oil fly ash (FOFA), or concentrated ambient particles collected from urban atmospheres. Exposure of animals to ROFA leads to inflammation, airway hyperresponsiveness, and lung injury (Dreher et al., 1997; Gavett et al., 1997; Killingsworth et al., 1997; Kodavanti et al., 1997b). Transition metals in ROFA appear to be responsible for many of these effects. Removal of metal from the fly ash attenuates the effects almost completely, and a leachate containing soluble metals provokes the response without insoluble PM constituents. In addition, a mixture of transition metals (Fe, Ni, V) can mimic many of the effects of ROFA. Different forms of ROFA or FOFA contain different relative and absolute concentrations of transition metals, and it is thus still difficult to identify which metals (or combinations of metals) are responsible for specific effects (e.g., epithelial cell damage, inflammation, hyperresponsiveness, etc.) from a mixture tested in vivo. Additional studies have provided evidence to suggest that bioavailable metals may be responsible for the acute pulmonary injury induced by ambient PM (Costa and Dreher, 1997).

Other PM components can play a role in making such metals bioavailable. Many combustion related particles contain "humic-like substances (HLS)," organic materials that are products of incomplete combustion and which can chelate transition metals and may be involved in the

production of reactive oxygen species (Ghio et al., 1996). The content of these substances is correlated with the concentration of acid-soluble transition metals in PM. Acid sulfates may react in emissions sources or in the atmosphere to form more available forms of transition metals. Soluble metals are often also present in PM as metal sulfates.

The potential for reactive oxygen species (e.g., hydroxyl radical) to cause or contribute to damage to lung cells was discussed in the 1996 CD. The in vitro cytotoxic effects of ROFA on rat tracheal epithelial cells was found to be due to the generation of reactive oxygen species (Dye et al, 1997). Iron and ultrafine titanium oxide have also been shown to lead to DNA damage via reactive oxygen species (Donaldson et al., 1996; Gilmour et al., 1996).

Combustion source PM-mediated airway inflammatory and fibrotic responses depend, in many cases, upon the metal components in the fly ash particles. Toxicity testing of individual metals have shown that nickel has a strong pro-inflammatory effect in vivo (Kodavanti et al., 1997a), resulting in PMN recruitment and airway hyperresponsiveness, and that vanadium, in epithelial cells in vitro, causes expression of pro-inflammatory cytokines (IL-6, IL-8, PGE2, PGF2α) (apparently through regulation of messenger RNA synthesis) (Samet et al., 1996; 1997; Carter et al. 1997). In vitro effects on macrophages, inflammatory mediator expression, DNA damage, and prostaglandin synthesis, all appear to be promoted by the metal component (especially the ionizable metal content (Pritchard et al., 1996) of fly ash particles.

Instillation of iron oxide particles (2.6 μm) in human lungs leads to a transient inflammatory response (Lay et al., 1997), possibly as a result of contamination of the iron oxide particles with intermediate iron compounds, capable of participating in electron transfer (Fenton) reactions. In the slow clearance phase (halftime >90 days), redistribution of particles from heavily laden alveolar macrophages to minimally loaded macrophages appears to occur. Becker et al. (1996) have shown that alveolar macrophages respond very differently to different types of particles than do epithelial cells, illustrating the importance of examining more than a single cell type from a single animal species. ROFA instillation also promotes synthesis of fibronectin from epithelial cells and macrophages, possibly via a stimulus which involves the production of reactive oxygen species. Fibronectin, is associated with repair of lung epithelial injury and hyperplasia, stimulation of fibroblast growth, and can be involved in the development of fibrotic lesions (Su et al., 1995). The extent to which fibrotic lesions develop may depend on the genetic makeup of the animal (or human) (Kodavanti et al, 1997a).

Because older individuals with chronic respiratory and cardiovascular disease appear to be disproportionately affected by PM exposure, many recent investigations have utilized animals that have been treated to mimic certain human diseases (so-called animal models of human disease) in order to investigate the effects of inhaled particles on chronic disease populations. For example, monocrotaline-treated rats (a model of pulmonary hypertension) exposed to ROFA also exhibit electrocardiographic abnormalities which lead to a higher mortality rate among these rats (Watkinson et al., 1997). Although the specific cause of death associated with the ECG abnormalities could not be identified, the ECG changes were consistent with progressive hypoxemic effects as well as acute electrical conduction disorders. FOFA-exposed monocrotaline-treated rats showed some evidence of inflammatory processes in the blood vessels and myocardium (Killingsworth et al. (1997). Using a device that excluded both ultrafine and

coarse particles, Godleski et al., 1997 exposed monocrotaline treated rats to concentrated ambient fine particles (CAPS) and also found increased mortality in monocrotaline treated rats. In the lung disease model rats (monocrotaline or chronic SO_2 exposure), ROFA, FOFA, and CAPS all caused increased mortality (Gavett et al, 1997; Godleski et al., 1997; Killingsworth et al. (1997; Watkinson et al., 1997). In the case of CAPS (Godleski et al., 1996), mortality increased 37% after exposure to approximately 270 $\mu g/m^3$ of CAPS (<2.5 μm) for 6h/day on three consecutive days in "SO_2-induced chronic bronchitic" rats. Although these studies add to the plausibility of epidemiologic PM-mortality findings, further work is necessary to establish the mechanisms of inhaled PM-induced cardiovascular effects.

Another possible mechanism for cardiovascular effects is illustrated in a recent study where phagocytosis, by macrophages, of deposited PM_{10} carbon particles in the lung leads to recruitment of PMN's from the bone marrow (Terashima et al., 1997). These new PMN's, although only a small part of the overall PMN recruitment, have a tendency to become "stuck" in the microvasculature where they produce enzymes, reactive oxygen species, and express L-selectin, an adhesion molecule, which may be associated with ischemic injury both in the lung and myocardium.

Other potential mechanisms for particle effects include allergic responses to (biological) antigenic substances in PM. The possibility that latex particles from automobile tire wear can cause allergic reactions in sensitized individuals has been examined (Miguel et al., 1996). Latex allergens in ambient air are present in quantities that are one thousandth (10^{-3}) of the amount necessary to provoke a response in a sensitive individual and appear to be less allergenic (presumably due to heat treatment in the formation of tires). Thus it is unlikely that PM responses can be attributed to allergic responses to latex antigens. A new study has examined the possibility that diesel exhaust carbon particles may carry pollen allergens (not pollen grains) on their surface (Knox et al., 1997; Frew and Salvi, 1997). They determined that diesel carbon can form aggregates with pollen acidic glycoproteins and thus facilitate the entry of micron-size pollen allergen fragments into the lung.

Churg and Brauer (1997) studied the size and composition of particles in lung tissue from 10 elderly persons and found that 96% of the particles were smaller than 2.5 μm in diameter and only a small portion (4.8%) had diameters of less than 0.1 μm (ultrafine particles). The authors hypothesize that the reason for the small numbers of ultrafine particles found is that inhaled ultrafine particles are phagocytized by macrophages and rapidly removed from the lung, and/or dissolved in lung tissue. Three new human studies (Bennett et al, 1997; Kim et al., 1996; Kim and Kang, 1997) demonstrate that individuals with chronic lung diseases (COPD, Asthma, "small airways disease") have a substantially increased deposition fraction of PM. Increased deposition fraction and increased ventilation rates at rest are estimated to lead to a two to three-fold increase in total particle deposition in patients with COPD who are exposed to $PM_{2.5}$. Some of the new information is useful in improving PM deposition prediction based on mathematical models. The deposition fraction (DF) factors for different respiratory diseases can be improved

based on the data of Bennett et al. (1997) and Kim et al. (1996, 1997). Bennett et al., (1996) also found no difference in DF between young (18-40 yrs) and old (60-80 yrs) and male and female healthy subjects exposed to 2 μm particles. Higher ventilation in males resulted in greater overall deposition. Theoretical (i.e. mathematical model) studies also suggest that the deposition in the nasopharyngeal and tracheobronchial airways of hygroscopic PM in the 0.2 to 3.0 μm range, based on consideration of dynamic humidity profiles, may be higher than previously estimated (Kaufman et al., 1996).

Many of these recently published studies provide new insight into possible mechanisms of action of inhaled particulate matter. The finding that inhaled PM can cause increased mortality in animal models of disease at $PM_{2.5}$ concentrations that exceed current peak ambient concentrations by less than a factor of three is particularly noteworthy in light of the epidemiologic observations of PM associated mortality. These newly published studies also suggest that further attention should be directed at the metal component of ambient PM. Sulfates may also play a role in the toxic effects indicated for soluble metals in PM, which is coherent with epidemiological observations that acid- or sulfate-bearing particles are associated with adverse health effects. Although many of the new animal toxicology studies involve instillation, not inhalation, of previously collected particles, this technique is appropriate to study mechanisms of action. Examination of other potential mechanisms (e.g., allergenic or size specific effects) is needed. Ongoing work on the effects of lung inflammation and PM phagocytosis in the lung on subsequent systemic effects, especially cardiac or vascular effects, is needed to provide further information on the relationship between inhaled pollutants and adverse cardiac events.

APPENDIX D. RESPONSE TO COMMENTS ON EXPOSURE MISCLASSIFICATION, THE ROLE OF INDOOR AIR POLLUTION, THE RELATIONSHIP OF AMBIENT AIR POLLUTION TO PERSONAL EXPOSURES, AND MEASUREMENT ERROR

A large body of commenters expressed concerns about the reliability of the epidemiological results because some studies showed a lack of correlation in cross-sectional comparisons between outdoor PM measured at central locations and indoor or personal exposures to PM (which includes PM from the outdoor, indoor and personal environments). By contrast, a few other commenters have argued that the PM results are confounded because the weather and other factors that cause daily variations in outdoor pollution will cause similar daily variations in indoor generated air pollution; in essence this means that personal exposure to indoor pollution would have to be correlated with outdoor pollution. These somewhat contrasting comments are addressed in Section I of this Appendix.

Commenters have also restated theoretically based concerns on a related issue, namely errors in the measurement of the concentrations of air pollutants. Some commenters have suggested that measurement error or, more generally, exposure misclassification, may have biased the effects estimates of PM, relative to co-occurring pollutants, which they argue would introduce further uncertainties in the estimated concentration-response relationships for all pollutants. EPA has conducted a more extensive examination of this issue, which is summarized in Section II.

I. RELATION OF AMBIENT PM TO PERSONAL OR INDOOR PM.

A. Introduction

A number of public comments addressed the relationship between ambient concentrations at a central monitoring site and personal exposure (i.e. the concentration at the breathing zone encountered by a person during his or her daily activities, both outdoors and in a variety of indoor micro-environments). This issue was addressed in great detail by Kenneth Brown in an article prepared for the American Petroleum Institute (API IV-D-2247) and in less detail by other commenters, including representatives of the Electric Power Research Institute (API IV-D-2329) and the Utility Air Regulatory Group (API IV-D-2250).

Dr. Brown's report "examines the relationship between concentrations of particulate matter measured by monitors worn by individuals (personal PM), and located in individuals' homes (indoor PM), nearby outdoors (outdoor PM), and central locations (central site PM)." Dr. Brown concluded that "The relationship between personal exposure to PM and outdoor PM levels varied from undetectable to weak or marginal, due to high intra-personal and inter-personal variability within and between cohorts. The contributions of indoor sources (emphasis added) to personal exposure was sizable compared to the outdoor component and also varied markedly among individuals." EPA agrees with these conclusions, which are in agreement with the

discussion in Chapter 7, Human Exposure to Particulate Matter: Relations to Ambient and Indoor Concentration, in the PM CD (U.S. EPA, 1996a).

Dr. Brown goes on to state, "The ultimate question is whether central monitoring site measurements contain a signal reflecting actual exposure at the individual level." Here Dr. Brown confuses the issue. The ultimate question is not whether the central monitoring site measurements contain a signal reflecting actual exposure to total PM from all sources at the individual level. Rather, the relevant question for this review is whether the central monitoring site measurements contain a signal reflecting actual exposure to ambient PM at the individual or population level, including both ambient PM outdoors and ambient PM that has infiltrated indoors. The key issue, then, is: can personal exposure to PM of ambient origin be adequately characterized by a central ambient monitoring station? This issue can be addressed most usefully in two steps. First, is the PM distribution over the area under study (usually a large city) sufficiently uniform so that a measurement at one ambient monitoring site (or the average of several) will provide an accurate average community concentration? Second, can the personal exposure to PM of ambient origin be adequately characterized by the outdoor concentration?

B. The Use of Ambient Monitors as an Index of Exposure to Ambient Air Pollution.

In areas with strong local point sources, the distribution of PM across the community on a 24-hour average basis would not be expected to be uniform. However, in most communities in the U.S. strong local sources have been relocated or controlled. In most communities in the U.S., and to some extent in other developed countries, PM pollution is largely due to a combination of regional pollution and a large number of smaller, more or less uniformly distributed, line, area, and point sources. In this case, the distribution of PM throughout the community, on a 24-hour basis, might well be reasonably uniform. According to the concept of Berkson error (see Section II below), it is not necessary that every person have the same personal exposure, only that the measurement used to characterize the subjects' exposure be an accurate measure of the average of the population's exposure to PM of ambient origin.

There are only a few areas and time periods for which measurements at several locations have been made over extended time periods. When seeking to determine an average community concentration, it is necessary to use some care in selecting monitoring data. Many monitoring sites were established to monitor major local sources and are located close to the source. Such measurements would not be expected to provide a meaningful characterization of community exposure. However, when analyses have been made of the site-to-site correlations of multiple sites (selected to measure community concentrations), the following has been observed (Wilson and Suh, 1997; Suggs and Burton, 1983):

(1) Outdoor (ambient) $PM_{2.5}$, PM_{10}, and TSP were found to be reasonably uniform across metropolitan areas (e.g., average correlation coefficient between Philadelphia sites, r = 0.89 for $PM_{2.5}$, 0.86 for PM_{10}, and 0.88 for TSP).

(2) Ambient $PM_{(10-2.5)}$ does not appear to be as uniform as $PM_{2.5}$, PM_{10}, or TSP, as shown in Philadelphia. However, it is not clear whether this is due to real variations in $PM_{(10-2.5)}$ across the area or to the lower precision of $PM_{(10-2.5)}$ compared to $PM_{2.5}$, PM_{10}, and TSP.

(3) For local-source oriented sites, in the same community, $PM_{2.5}$ was reasonably uniform. However, $PM_{(15-2.5)}$, PM_{15}, and TSP were not.

The PTEAM study measured PM_{10} concentrations at 178 homes (day and night) over a 45 day period in Riverside, California. (Özkaynak *et al*., 1996b). During the study period, PM_{10} was also measured at a central site. A correlation coefficient, r = 0.75 was found between the central monitors and the backyard PM concentrations (Figure D-1). (If 5 clear outliers were removed, a much better correlation would have been obtained.) Other studies have observed that indicators of fine-mode particles such as sulfate and acidity are uniform over both community and regional areas (Keeler *et al*., 1991; Lee *et al*., 1993; Liu, *et al*., 1996; Suh, *et al*., 1995).
It would be desirable to demonstrate uniformity of distribution before using a PM indicator as a measure of exposure in epidemiological studies. However, in those cases in which it has been tested in the U.S., concentrations of $PM_{2.5}$, PM_{10}, and TSP, as measured by community monitoring sites, are adequately uniform across the area of interest to provide an average value of community concentration.

It is true that personal exposure to total PM (PM from indoor sources and outdoor sources) is not well characterized by outdoor concentrations, however, this statement is of limited relevance here. What is true and relevant, as stated clearly in Chapter 7 of the PM CD (U.S. EPA, 1996a), is that personal exposure to **PM of ambient origin** (PM from only outdoor sources, experienced indoors as well as outdoors) is indeed well characterized by outdoor concentrations. The difference between Total Personal Exposure to PM and Personal Exposure to Ambient PM is illustrated in Figure D-2.

Lipfert and Urch state that "indoor RSP [respirable particles of undefined size] was a significant predictor of personal RSP ... This relationship assigns a very minor role to RSP from outside sources". They neglect the fact, as shown in Figure D-2, that indoor RSP is composed of PM from indoor and outdoor sources.

Misunderstanding of relationships between ambient PM measurements and personal exposures arise from a paradox derived from cross-sectional analyses in which different people are pooled together from different homes as though their measurments were independent. The simple example presented in Figure D-3 (U.S. EPA, 1996a) shows how pooling the personal exposures of two subjects, each with total personal exposure perfectly correlated with ambient PM data (R^2 = 100%), creates an appearance of an imperfect correlation (R^2 = 7%). However, the average

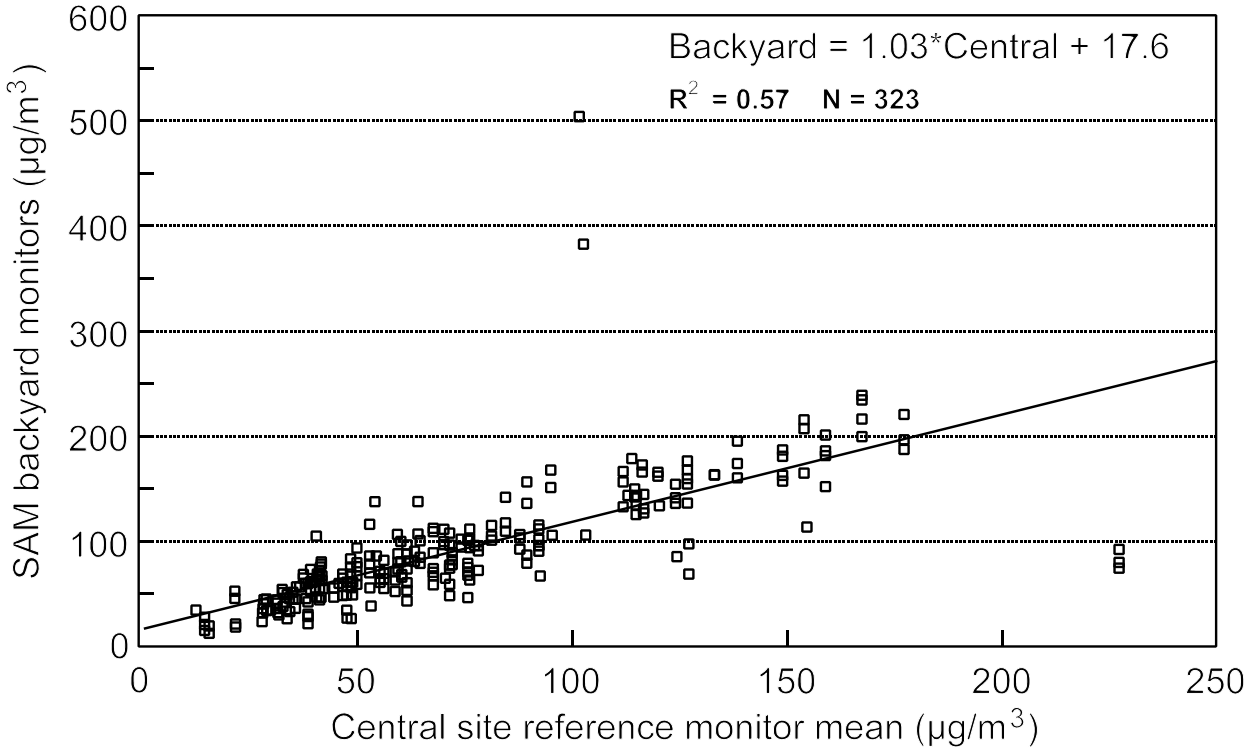


Figure D-1. Residential outdoor monitors versus central-site mean of two dichotomous samplers in Riverside, CA, R^2 = 57%. Source of Data: Pellizzari *et al*. (1993) . (After U.S. EPA, 1996a; Figure 7-20).

exposure of the two subjects falls on the regression line and has a perfect correlation (R^2 = 100%). Figure D-4 shows how this relationship holds in practice. Personal exposure data for PM_{10} of 14 subjects in Phillipsburg, NJ, vs ambient PM_{10} data on the same days, have an R^2 = 25% when the individual observations are pooled together. When the observations on a given day are averaged together, the regression line between average personal exposure to PM_{10} and average ambient concentration of PM_{10} is R^2 = 90% (U.S. EPA, 1996a). Although some commenters (e.g., AISI, IV-D-2242) noted that the PTEAM study reported poor serial correlations for both PM_{10} and $PM_{2.5}$, Table 7-6 of the CD showed that "The R^2 values improved considerably when the regressions for individual homes were calculated" and many individual exposures had highly significant correlations with the ambient PM data. Moreover, data from Tamura *et al*., (1996), presented in the CD, show clearly how the serial correlations between personal exposure to total PM and ambient PM are strong and statistically significant when the sources of indoor PM generation are minimized by examining people in homes with minimal sources of PM and

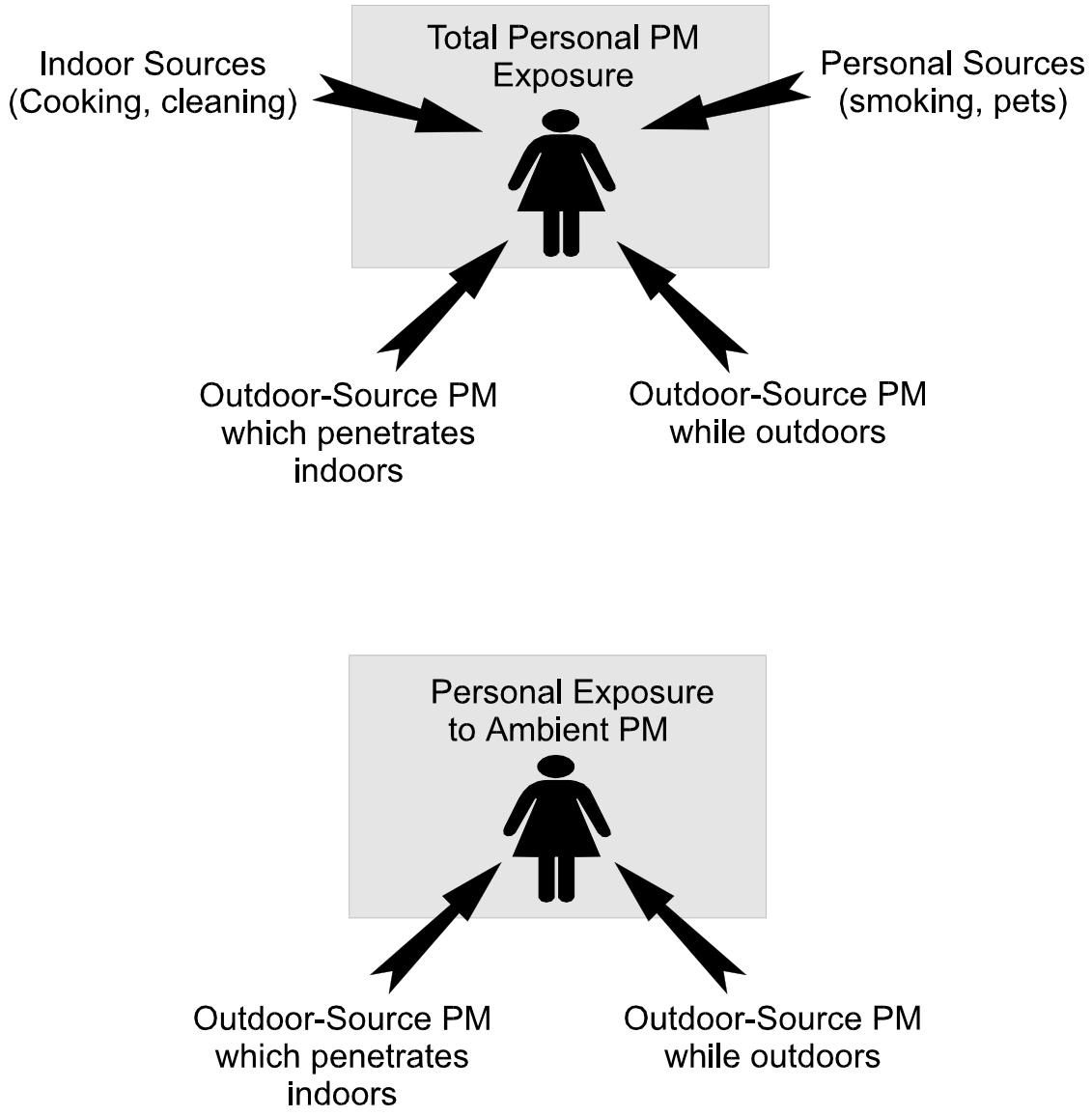


Figure D-2. Difference between Total Personal PM Exposure and Personal Exposure to Ambient PM.

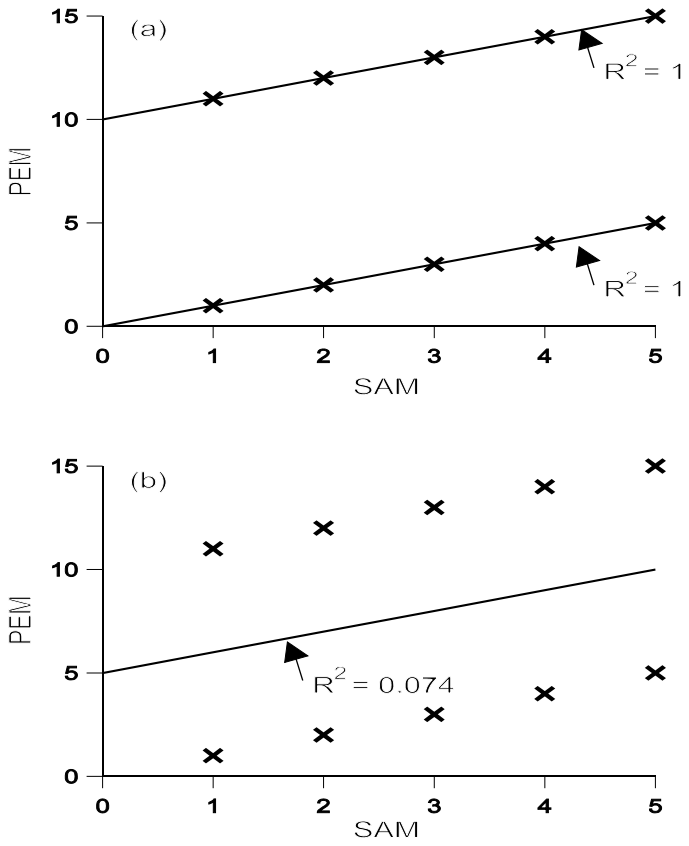


Figure D-3. Example of difference between serial correlation (a) and cross-sectional correlation (b) of personal exposure monitoring (PEM) and stationary ambient monitoring (SAM), showing how pooling of individuals together can mask an underlying relationship of PEM and SAM. (After U.S. EPA, 1996a; Figure 7-28)

- Let two people live next door to each other at a location where the ambient PM for 5 consecutive days has a sequence {1, 2, 3, 4, 5}.

- Let person A live without environmental tobacco smoke (ETS) exposure and have a corresponding PEM series {1, 2, 3, 4, 5}, ($R^2 = 1$).

- Let neighbor B live with ETS exposure and have a corresponding PEM series {11, 12, 13, 14, 15}, ($R^2 = 1$).

- When their PEM values are pooled so that they are analyzed together (cross-sectionally) {(1,11), (2,12), (3,13), (4,14), (5,15)} vs the SAM set {1, 2, 3, 4, 5}, then $R^2 = 0.074$.

- However, had the two PEM series been averaged each day, the sequence of averages {6, 7, 8, 9, 10} would have a correlation of R2 = 1 with the same SAM sequence. This averaging process is described in more detail in CD Section 7.6.2.

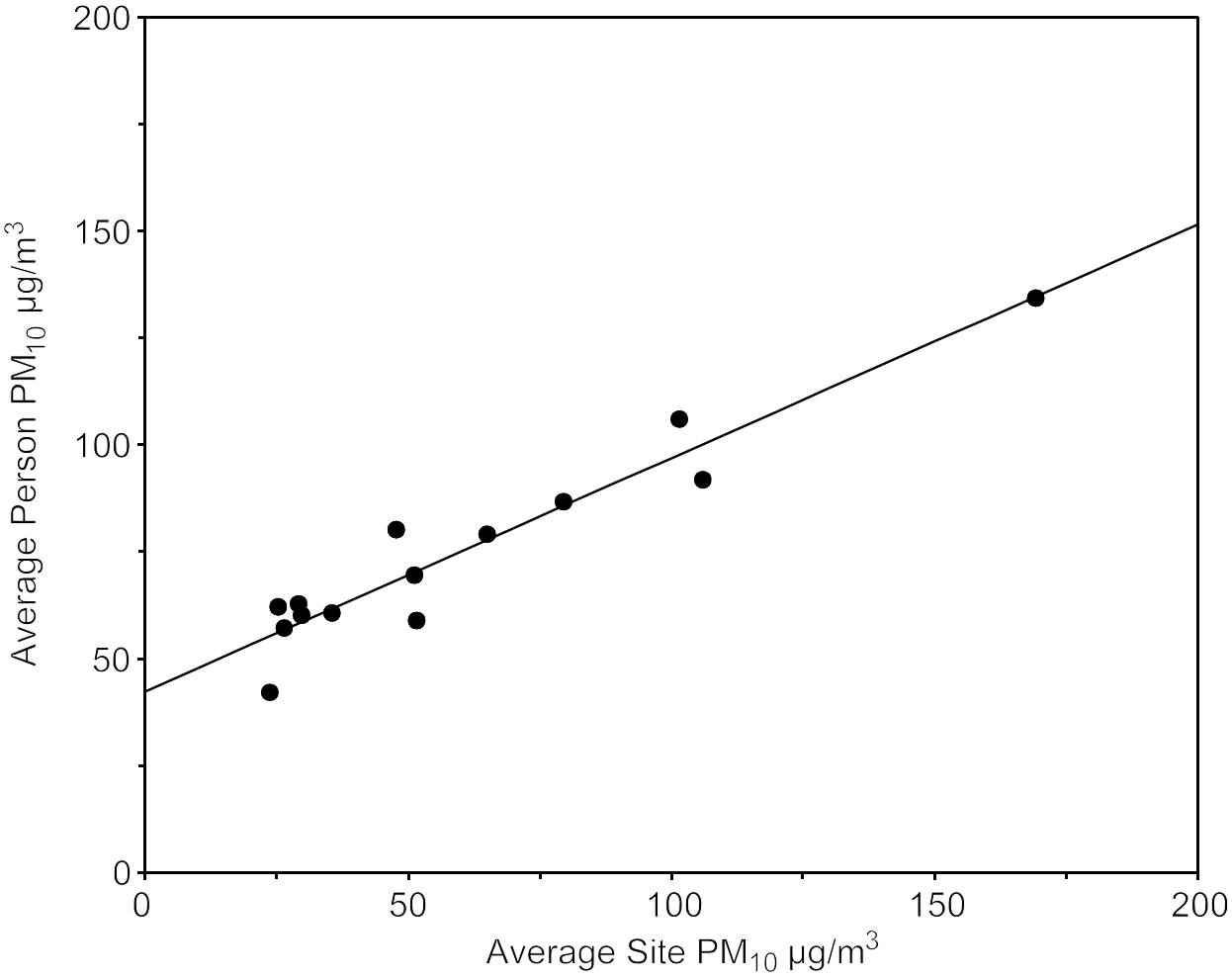


Figure D-4. Plot of relationship between average daily personal PM_{10} exposure versus ambient PM_{10} monitoring data from Phillipsburg, NJ and regression line calculated by U.S. EPA based on Lioy et al. (1990). (After U.S. EPA, 1996a; Figure 7-32)

culling out the data when smokers and other combustion sources were operating in the home. Indoor PM_{10} is strongly correlated with the outdoor PM_{10} for the seven homes in this study (see CD Figure 7-24). Although one commenter (Wyzga IV-D-2672) claimed that the Tamura *et al*. studies in Japanese homes are different from U.S. homes (tatami flooring vs wall-to-wall carpeting), similar high correlations would also logically be expected in U.S. homes where indoor sources of PM are minimal. Thus the Tamura study provides evidence for the strength of relationship between ambient PM concentrations and personal exposure to PM of ambient origin.

C. Analysis of Relationships Between Personal Exposure to PM of Outdoor Origin and Outdoor PM Concentration.

This section shows how personal exposure to ambient PM (Ea,p) can be estimated and then analyzes data from the PTEAM study. This section derives a statistical relationship between personal exposure to PM of outdoor origin and outdoor PM concentrations.

Theoretical Relationships. The prediction of the relationship between personal exposure to **PM_{10} of ambient origin** and ambient PM_{10} data was presented by the CD as an equation derived as follows:

Let $Ep = Ea + Ein$, (1)

where Ep = Total personal exposure to PM_{10},

Ea = Exposure to PM_{10} of ambient origin while outdoors, assuming that indoor generated PM_{10} has no perceptible influence on ambient concentrations of PM_{10},

Ein = Exposure to PM_{10} of ambient and indoor origin while indoors.

Then $Ea = y\ Ca$, and (2)

$Ein = [Ca\ P\ a/(a + k) + Cin]\ (1 - y)$, (3)

where Ca is concentration of PM_{10} of ambient origin measured outside the home,

Cin is concentration of PM_{10} of indoor origin in the subject's home,

P is fraction of PM in the ambient air that penetrates into a home ($P = 1$ means no losses)

a is the air exchange rate measured in the subjects home, hr^{-1},

k is the deposition parameter, hr^{-1}, a function of size distribution of the PM, and includes an assumption that there is no auxiliary filtration of indoor air,

y is fraction of time spent outside the home (outdoors, in car, at work, ...).

Then, $Ep = [(1 - y)\ P\ a/(a + k) + y]\ Ca + Cin\ (1 - y)$, and (4)

$Ep,a = [(1 - y)\ P\ a/(a + k) + y]\ Ca = z\ Ca$ (5)

where Ep,a = The total personal exposure to PM_{10} of ambient origin, and

$z = [(1 - y)\ P\ a/(a + k) + y]$, the fraction of ambient concentration of PM_{10} to which the subject is exposed.

Equation 5 is analyzed for the conditions of the Riverside, California, PTEAM study, using the measured daytime data for the PM_{10} backyard concentration for each subject (Ca), the measured daytime air exchange rate for each subject's residence (a, including replacing values when air exchange rate exceeded the maximum detectable level (MDL) with the MDL), each subject's recorded fraction of time spent out of the home (y) during the day, and results derived from a mass-balance model that indicate that the home penetration factor (P) for PM_{10} is virtually unity and the average deposition parameter (k) is 0.91/hr as estimated for this cohort by Özkaynak *et al*. (1996 a,b). Riverside, California, is a typical urban community, with no major local point sources of PM. The major sources of PM are those associated with the ubiquitous motor vehicle traffic and urban activity of the Los Angeles Metropolitan area, so the results are not atypical of such urban areas.

Two non-independent estimates can be made for *z* using the PTEAM data set:

(1) for the average daytime value of a = 1.08/hr, average daytime y = 0.211, and k = 0.91/hr as a daytime average value, one obtains a single value for z = 0.639, and;

(2) for the average value of *z* found by using the average value of *k* and the individual values of *y* and *a* for the same subjects, one obtains a population average of z = 0.562.

The daytime value of k (0.91/hr) for PM_{10} is much greater than the nighttime value of *k* (0.43/hr), perhaps because of the decrease in size of the suspended PM in the home when human activity is not resuspending PM or generating PM while the subjects are sleeping. A lower value of k than 0.91/hr during the day would therefore be applicable to the ambient PM fraction that entered the home. If the average k value for both night and day of 0.65/hr is used to represent the value for the daytime ambient PM fraction, then z = 0.647.

Experimental Verification. We compare these theoretical estimates with the regression of daytime measured personal exposure (Ep) to PM_{10} vs the individual's backyard PM_{10} concentration measurement (Ca). Figure D-5 shows that Ep vs Ca has a regression line with a slope of 0.72 and a regression coefficient of R^2 = 15%. This slope means that, on the average, for each unit increase in ambient PM_{10}, the subjects' average exposure went up by 72% of that unit increment.

Özkaynak and Spengler (1996) report additional analyses of PTEAM data. Figure D-6 shows a combined set of daytime and nighttime PTEAM data (N = 294) on Ep vs the respective individual day and night ambient PM values (Ca), with a slope of 0.61. Figure D-7 shows these same values averaged over different subjects (typically three per day) for a given nighttime and daytime period (N=94), with a slope of 0.66.

Discussion. The excellent agreement between theory (56%, 64%, 65%) and observation (61%, 66%, 72%) is a confirmation of the theoretical model described above, even considering all the various simplifying assumptions that went into the derivation used in developing the equations 4 and 5 (e.g., when the subjects leave home they spend time in locations either outdoors or indoors, as in a car, with very high air exchange rates and minimal sources of indoor PM_{10}; the k value is equal to 0.91/hr during the day for each subject; etc.). The low values of R^2, 15% in Figure D-5 and 16% in Figure D-6, are artifacts of combining people with similar exposures to ambient PM_{10} who also have widely different sources of PM in their homes, as explained by the example situation shown for Figure D-3. Figure D-7 shows how averaging personal exposures of different individuals reduces the noise introduced by non-ambient indoor PM_{10} sources, and that it increases the R^2 value from 15% to 33% (also see Figure D-4).

In conclusion, this analysis shows that people are exposed to PM of ambient origin, even if they remain indoors, and that variations of ambient concentrations of PM are reflected in variations of personal exposure to that ambient PM. This establishes the vital linkage between the ambient concentrations of PM which are influenced by the sources controlled by EPA, and the personal

exposure to these particles of ambient origin that are hypothesized to create the health effects that have been discovered through the peer-reviewed epidemiological investigations reported in the CD.

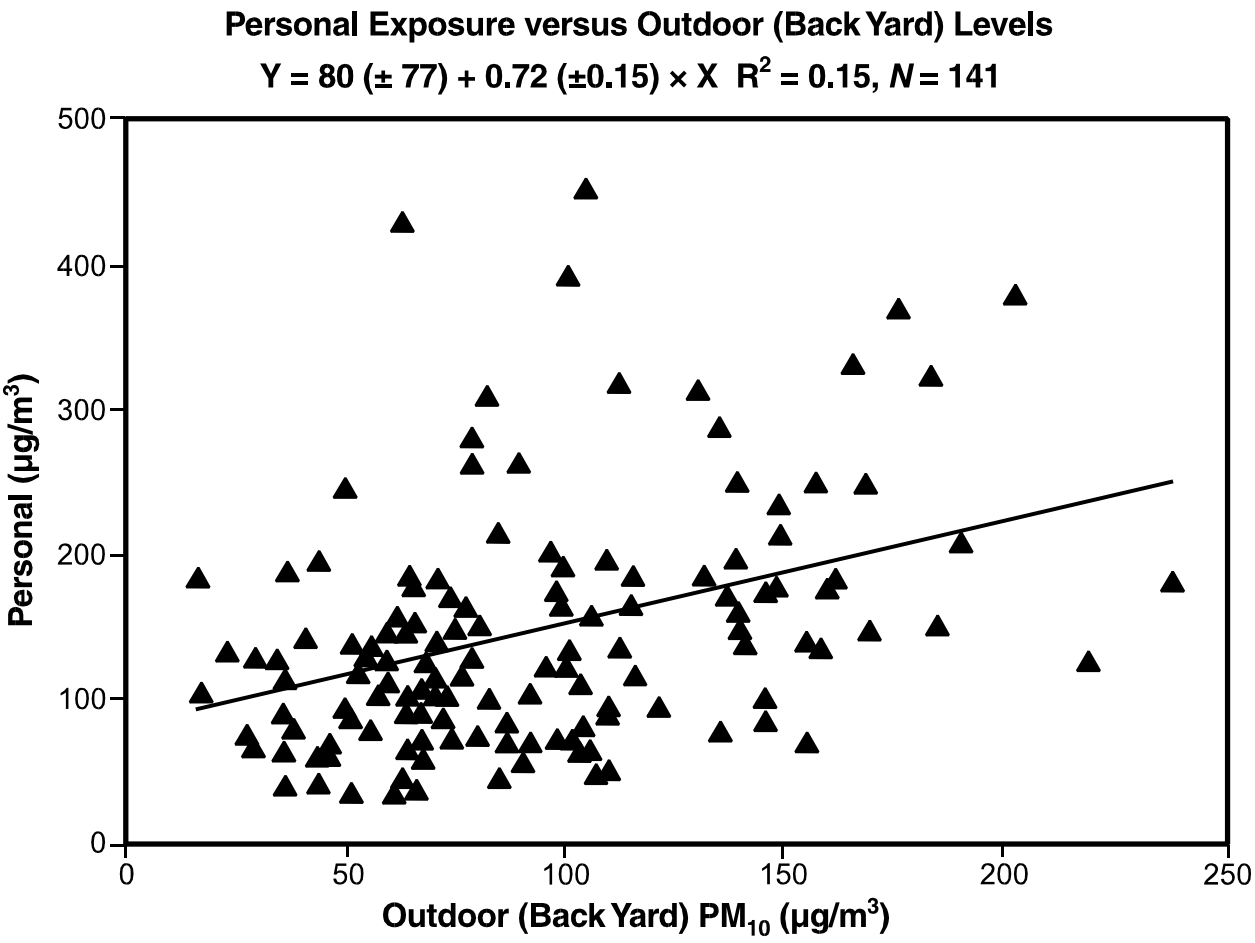


Figure D-5. Personal exposure to PM_{10} of Riverside, CA, residents during the daytime period versus the concentration of PM_{10} measured simultaneously in their own back yard (N = 141, two high back yard values are omitted from the regression). The results show that for every 1.00 µg/m^3 increase in average outdoor concentration, the average personal exposure increases by 0.72 µg/m^3.

Source: PTEAM data (Ozkaynak et al., 1996a,b) analyzed by EPA for this document.

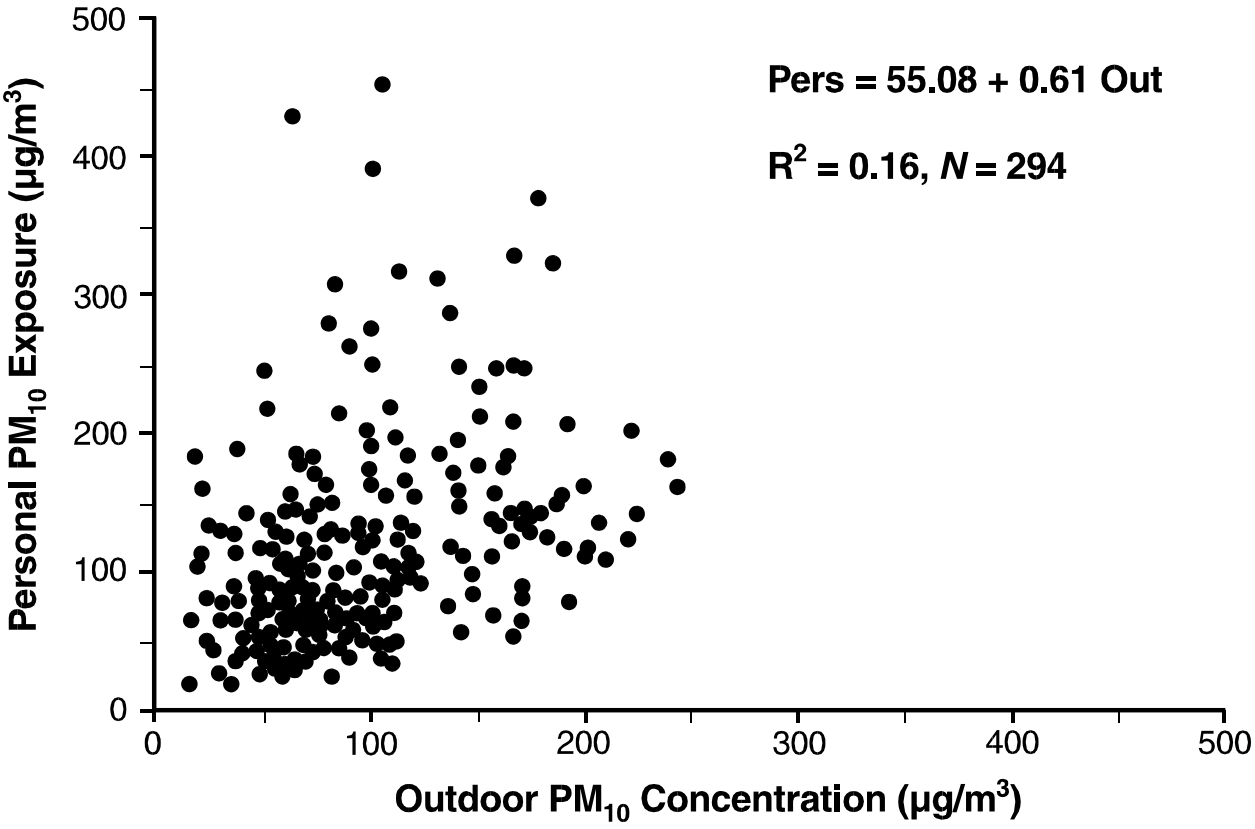
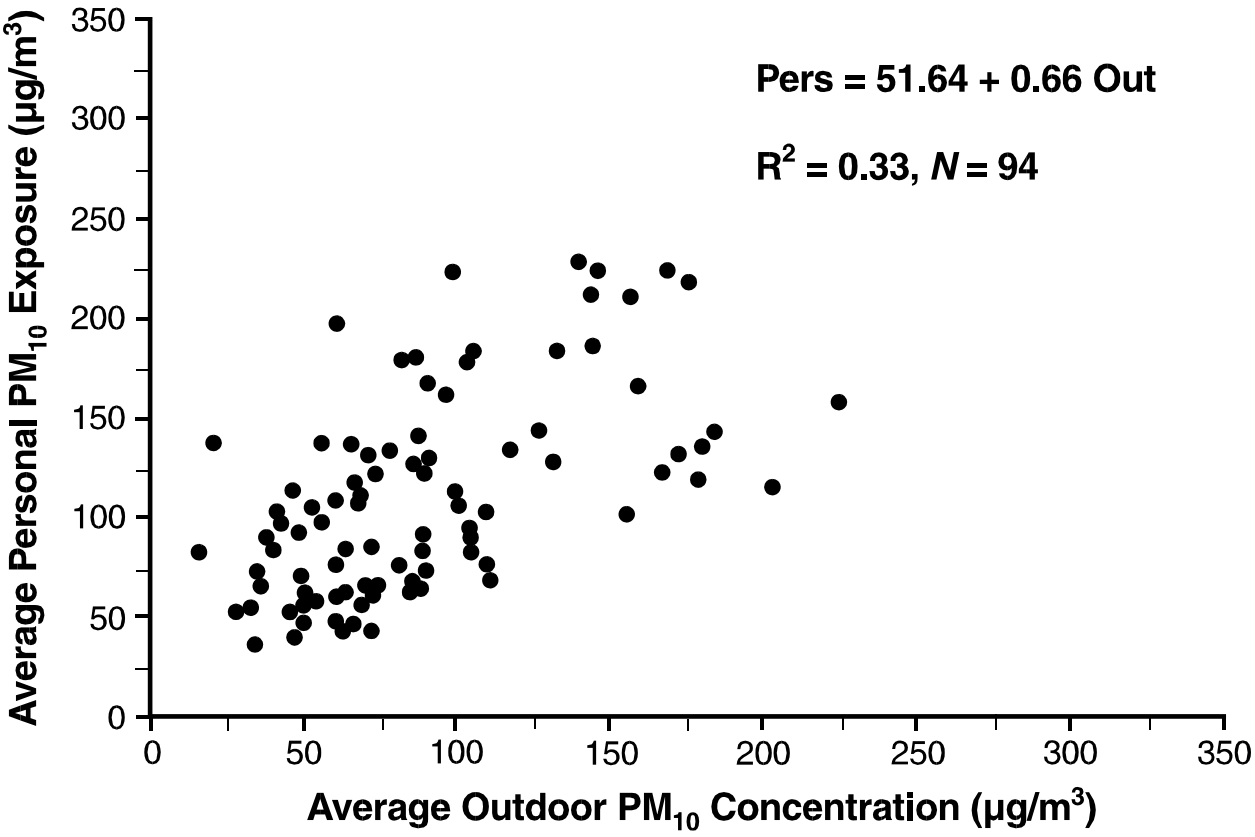


Figure D-6. Comparison of daily day time and night time individual personal and outdoor PM_{10} concentrations in Riverside, CA.

Source: PTEAM data (Ozkaynak et al., 1996a,b) analyzed by EPA for this document.



7

Figure D - 7. Comparison of daily averaged personal and outdoor PM_{10} concentrations in Riverside, CA.

Source: PTEAM data (Ozkaynak et al., 1996a,b) analyzed by EPA for this document.

D. Analysis of the relationship between indoor-generated and outdoor-generated PM.

Some commenters claim that the results of the epidemiological studies showing relationships with ambient PM concentration are confounded by indoor PM generated in the homes of susceptible individuals. To confound these results, concentrations of indoor-generated PM would need to be correlated with ambient PM concentration and also be responsible for the health effects that are correlated with ambient PM concentration. The following analysis shows that this situation is not likely.

Some commenters have hypothesized that people spend more time indoors on higher pollution days, so they may be exposed to more indoor PM on those days, so the indoor-generated PM might be the cause of the effects correlated with the ambient PM. The PTEAM data base reports fraction of time in-residence and backyard PM for 149 daytime cases where individuals had their personal exposures monitored. The correlation of fractions of time spent indoors at their residence (Y) vs the concentration in the backyard (Ca) showed that these quantities were virtually independent of each other:

$Y = 0.425 + 0.0009\ Ca \quad (r = 0.05;\ R^2 = 0.25\%)$.

This is an important relationship because PM values in Riverside, California, are amongst the highest in the U.S. Consequently, there is no reason to believe that the lower values in the cities where mortality studies were conducted would make the people spend more time indoors on the relatively higher local PM days, since those values would likely be below the values reported in Riverside.

The non-ambient source PM component of the personal exposure can be estimated by subtracting the ambient PM portion of the personal exposure ($Ep,a = z\ Ca$) from the personal exposure (Ep). Figure D-8 presents the regression of $Ep - Ep,a$ *vs* Ca which shows no significant increase in exposure to PM_{10} generated by indoor sources as ambient PM_{10} increases. This behavior is expected because people do not cook or smoke more on higher ambient pollution days than on lower ambient pollution days. Then, if both exposures to PM of ambient origin and exposures to PM of indoor origin cause health effects, these health effects must vary independently of each other because the fluctuations of indoor PM are driven by processes (e.g., vacuuming, dusting, smoking, cooking, hobbies) that are independent of, and uncorrelated with, the subjects' exposure to the total PM_{10} of ambient origin. Thus, fluctuations in exposure to indoor-generated PM cannot be a true confounder in the statistical analysis of the association between the fluctuations in health parameters that are correlated with the fluctuations of ambient PM.

Arguments such as those made by Vedal (1997) and those made by Lipfert and Wyzga (1997) are thereby shown to be based upon false premises and are largely refuted by data analyses presented, as follows:

Vedal (1997) states -- "if study subjects closed their windows on days with higher levels of pollution, exposure to indoor pollutants might increase and actually be responsible for the

increase in adverse health outcomes, and therefore confound the particle and health association." This argument is rejected by examination of the experimental PM_{10} data and air exchange data from PTEAM. A regression was run for the PTEAM data set of N = 149 matched values of daytime indoor PM_{10} (Cin) and the air exchange rate (a). The reciprocal ratio $1/(a + k)$, with k = 0.91/hr, representing the mean residence time of indoor generated PM_{10} in the home during the day, was used in a regression because this quantity is expected to be more linearly related to the concentration of indoor generated PM_{10} than the air exchange rate itself. The concentration of PM_{10} of ambient origin that is estimated in the residence [Ca $a / (a + k)$] was subtracted from the indoor PM_{10} concentration (Cin). This quantity, Cin - Ca $a / (a + k)$, showed no appreciable increase with residence time which would be associated with closing of windows. The regression results are as follows:

$$\text{Cin - Ca } a/(a+k) = 47.8\ \mu g/m^3 - 6.1\ \mu g/m^3\text{-hr}/(a+k) \qquad [r = -0.03, (r = 0.1\%)]$$

This result complements the previous finding that the indoor-generated PM appears to be independent of, and uncorrelated with, the ambient pollution measurements in the backyard of each home.

Lipfert and Wyzga (1997) state-- "We refer to the differences between the values obtained from ambient monitors and actual (total) personal exposures as 'exposure errors.' " They fail to make the necessary distinction between the personal exposure to PM from ambient sources and the actual (total) personal exposure to both ambient PM and indoor PM as collected by a personal monitor whilst a subject spends time indoors and outdoors. Since the indoor-source PM portion of personal exposure has been shown here to be independent of the ambient portion, the error they refer to here is not an error at all in the process of establishing whether or not there are any fluctuations of health endpoints correlated with fluctuations of ambient PM concentrations.

There is also an excellent relationship (R^2 = 82% in Riverside, CA) predicted from the PTEAM data, between personal exposure **to PM_{10} of ambient origin** (Ep,a computed by Equation 5) and ambient monitoring data for PM_{10} in the subject's back yard (Ca). This result is not an unexpected finding because Ca appears in both terms, but it needs to be experimentally verified for PM_{10} by means of measuring the amount of the collected PM_{10} by the personal monitor that originated from ambient sources, which is unavailable at present. However, this relationship applies directly for sulfates which have no appreciable sources indoors. Figures D-9 and D-10, from Suh *et al*. (1993), and Özkaynak *et al*. (1996), show that their reported sulfates, which are submicron-sized, had an R^2 value of 92% and 77% respectively, between personal exposure to sulfate and outdoor sulfate. For Figure D-9, the slope is 0.78, and for Figure D-10, the slope is 0.70 which supports the concept that the relationship between exposures to PM of ambient origin and ambient PM concentrations is driven by air exchange rates, indoor deposition parameters and fractions of time spent outdoors. Because the deposition parameter (k) for $PM_{2.5}$ is intermediate to that for PM_{10} and sulfate (PM <1) it is expected that the relationship between personal exposure to $PM_{2.5}$ and $PM_{2.5}$ of ambient origin would also be intermediate to those of PM_{10} and sulfates presented here.

It is important to note here that due to the more rapid removal of larger particles in indoor environments, the relationship between ambient concentrations and personal exposure to ambient PM will be stronger for accumulation mode particles than for TSP, PM_{10}, or to a much lesser extent, $PM_{2.5}$ since all of these measurements contain some coarse mode particles. The agreement will be poorest for indicators of coarse mode particles (such as $PM_{10-2.5}$). This results in a smaller difference between the indoor concentration of ambient PM and the outdoor concentration of ambient PM for the smaller fine particles than for the larger coarse mode thoracic particles. Along with the differences in chemistry and sources of fine and coarse mode particles, this supports EPA's decision to treat fine-mode and coarse mode particles as separate classes of pollutants.

The effect of air-conditioning (A/C) on indoor concentrations of ambient PM has been raised by some commenters (e.g. Ford Motor Co., IV-D-5323). The presence of an operating air conditioner may lead people to tighten their homes which would reduce the air exchange rate and reduce the equilibrium concentration of ambient particles which have infiltrated indoors. In addition, both indoor PM and outdoor PM which has infiltrated indoors may be filtered out of the return air to the A/C system by an in-line filter. One commenter claimed that use of the air conditioner may raise indoor PM concentrations, perhaps by resuspending PM. Although this resuspension is possible for coarse mode PM of both indoor and outdoor origin, fine PM is not resuspended by air currents because fine PM is more tightly bound to surfaces, as shown by the data sets for fine PM species with no indoor sources, such as sulfates. Figure D-9 shows that personal exposures to sulfates are lower when air conditioning is turned on. The effect of this A/C filtration is thus to lower the correlation between ambient sulfate and personal exposure to sulfate. However, for coarse mode PM of both indoor and outdoor origin there could be an opposite effect. The operation of an A/C system with the combined mechanical and convective circulation of cold air in at the inlet and warm air out at the return to the A/C system, could lead to resuspension of coarse PM. Rojas-Bracho *et al*. (1996), found "no [significant] correlation" between total personal PM_{10} and $PM_{2.5}$ exposures of non-smoking COPD patients in A/C homes and the simultaneous ambient PM concentrations measured outside their homes. A/C-created resuspension of coarse mode PM, to the extent it occurs, would drive the correlation between ambient PM and personal exposure PM towards a range not significantly different from zero. The circulating concentration of PM of ambient origin in an A/C home will still be positively correlated with the ambient PM concentration, as indicated by Figure D-9. This follows because while the A/C is in operation and indoor sources and air exchange rate are held constant, an increase in ambient PM must cause a proportionate increase in the indoor concentration of PM of ambient origin, albeit less than the increase that would have occurred if the A/C were not in operation and the home were open for ventilation.

E. Summary and Conclusions

In conclusion, current evidence strongly suggests that personal exposures to indoor-generated PM vary independently of the personal exposures to PM of ambient origin, and personal exposures to PM of ambient origin are expected to be a major portion of the ambient PM measured in a person's residential area. Because of the limited number of sources of fine PM in an indoor microenvironment when smokers are not present (i.e., fine PM is not resuspended from

surfaces by human activity) personal exposure to $PM_{2.5}$ will have an even higher correlation with the ambient $PM_{2.5}$ measured in a person's residential area. Therefore a valid measurement of ambient PM can be related by a constant factor of proportionality to the average exposure of the people in that community to PM of ambient origin, and that factor will increase as the particle size decreases; in Riverside, California, for the fall period of 1989, that constant factor for PM_{10} was approximately $z = 2/3$ so that this approach appears to be valid for such a community. Other communities may be expected to have different characteristic proportionality constants for the fraction of PM_{10} of ambient origin that residents are exposed to because the seasonal temperatures may be higher or lower, so windows may be more or less open than in Riverside, and homes might be more or less tightly sealed than in Riverside (resulting in a different distribution of air exchange *a*), and the people could spend more or less time out of doors (a different distribution of *y*) than for the Riverside cohort. For example, in Phillipsburg, New Jersey, the relationship of Figure D-4 is $z = 5/9$ which is less than the value of $z = 6/9$ for Riverside, which may be explained by tighter homes (lower air exchange rates) and more time spent indoors in the winter for the New Jersey subjects. However, in any community, although the regression equation slopes may differ, there will typically be a strong relationship between personal exposure to PM of ambient origin and ambient concentrations of PM.

Consequently, if there is a health effect associated with exposure to PM of ambient origin, (and that health effect is not associated with some other unknown pollutant that is highly correlated with ambient PM concentrations) then the ambient PM concentration measured in a community is proportional to the average exposure of people in that community to PM of ambient origin and it is a mathematically valid surrogate for the agent causing the health effects. Therefore, any such claims (e.g., Crandall *et al*., 1996; Gamble and Lewis, 1996; Wolfe, 1997; Vedal, 1997; Lipfert and Wyzga, 1997), that total personal exposure to PM is uncorrelated with ambient PM concentrations and that therefore exposure to indoor PM that is generated by indoor sources may be responsible for the health effects found to be associated with ambient PM fluctuations, appear to be due to a misunderstanding of the true relationship between exposure to PM of ambient origin and ambient concentrations of PM.

There is ample evidence, as discussed in Chapter 7 of the PM CD, that personal exposure to ambient PM, while outdoors and while in indoor micro-environments, does correlate with concentrations measured at central monitoring sites. However, since exposure to PM generated indoors or by personal activities is not correlated with central site PM concentrations, such exposures would not be a confounder in epidemiological studies. They could indeed be independent risk factors. Because personal exposure to ambient PM is not correlated with personal exposure to indoor or personally-generated PM, the health effects found to be associated with changes ambient PM in epidemiological studies are not attributable to indoor or personally-generated PM. The proposed PM standard is intended to protect the public from exposure to ambient PM. It is not intended to protect them from PM generated in their homes or by their personal activities.

Indeed no significant correlation has been found between ambient concentrations and personal exposure to PM from all sources, except in unusual situations. However, this information is irrelevant to EPA's regulation of ambient PM. The salient factor is that there is a relationship

between health outcomes and ambient PM concentrations and between ambient PM concentrations and personal or population exposure to ambient PM.

EPA has concluded that it is reasonable to presume that a reduction in ambient PM concentrations will reduce personal or population exposure to ambient PM concentrations and will thereby help to protect the public from adverse health outcomes associated with personal exposure to ambient PM. EPA has taken no position on the possible health effects of indoor- or personally-generated PM.

Current evidence strongly suggests that exposure to PM generated indoors, or by personal activities while outdoors (such as smoking), is not correlated with central site PM concentrations. Therefore, such exposures would probably not be a confounder in epidemiological studies. They could indeed be independent risk factors. Conversely, because personal exposure to ambient PM is not correlated with personal exposure to PM generated indoors or personally generated PM, epidemiological studies relating health outcomes to ambient PM would not provide any information about the health effects that may be caused by PM generated indoors or by personally generated PM.

EPA agrees that indeed there may not always be a significant correlation between ambient concentrations and personal exposure to total PM from all sources. However, this information is of limited relevance to EPA's regulation of ambient PM. The significant factors are that: (1) there is a clear relationship between health outcomes and ambient PM concentrations; and, (2) there is a clear relationship between ambient PM concentrations and personal exposure to ambient PM.

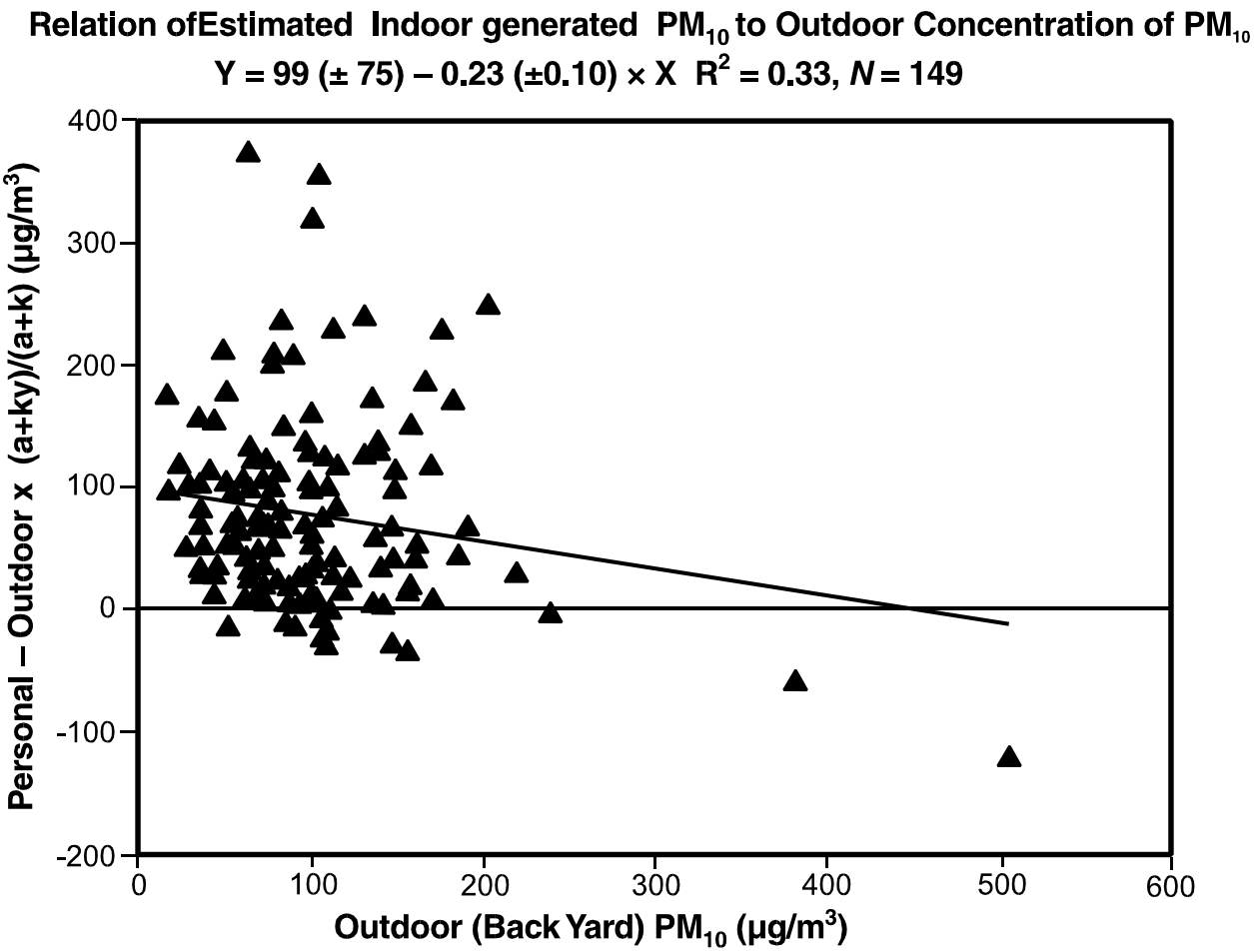


Figure D-8. Personal exposure to total PM_{10} minus the estimated personal exposure to PM_{10} of ambient origin is an estimate of personal exposure to PM_{10} of non-ambient origin as described by Figure D-2 This quantity is plotted above versus the simultaneous back yard concentrations for ambient PM_{10}, for residents of Riverside, CA, during the daytime hours (N = 149). The results show that residents of Riverside were not likely to have been exposed to more PM of non-ambient origin on days with higher ambient air pollution. This establishes that the 0.72 microgram personal increase per microgram ambient increase shown in Figure D-5 virtually comes entirely from the ambient PM sources.

Source: PTEAM data (Ozkaynak et al., 1996a,b) analyzed by EPA for this document.

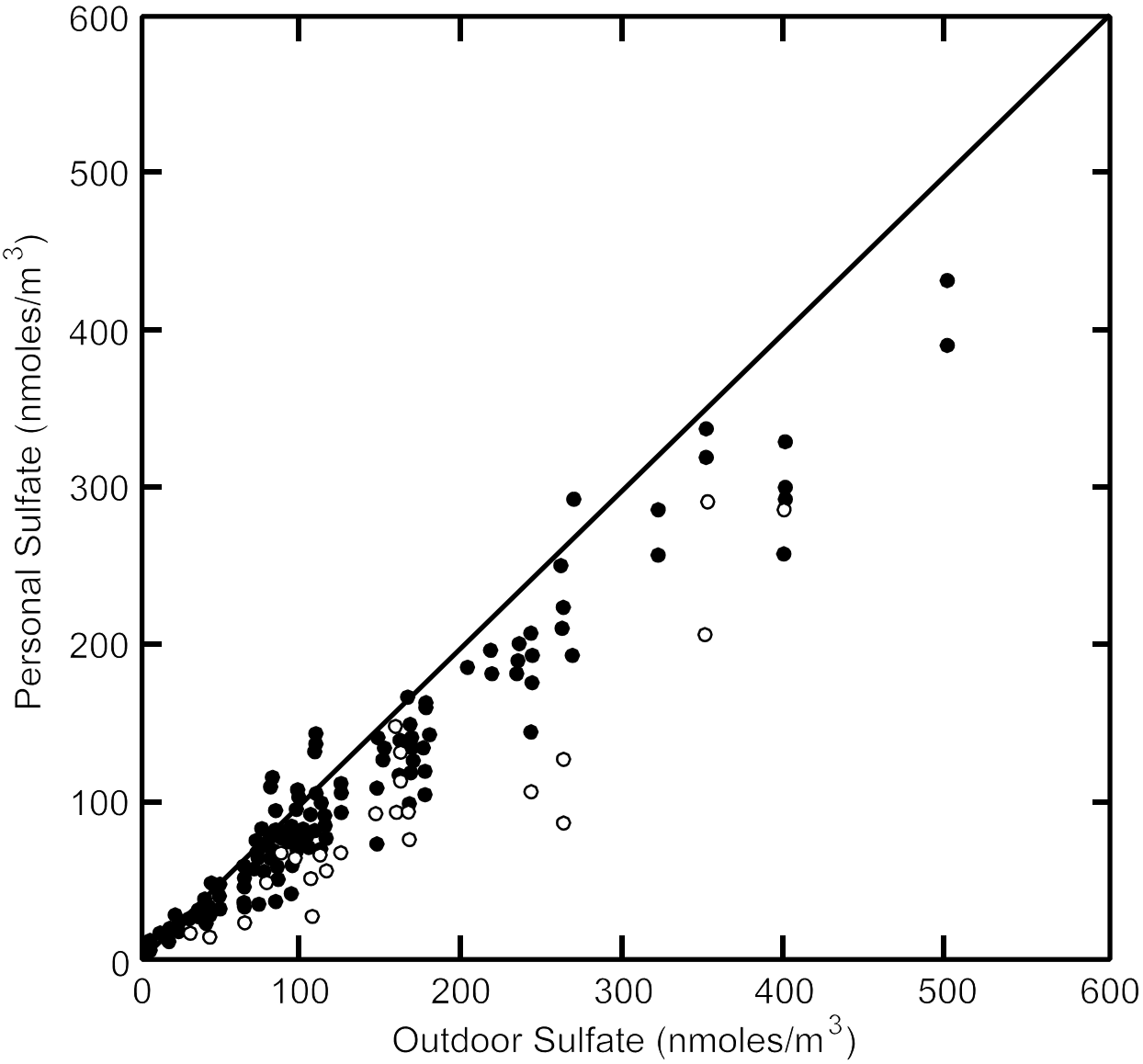


Figure D-9. Personal versus outdoor $SO_4^=$. Open circles represent children living in air conditioned homes; the solid line is the 1:1 line. These results show the strong relationship between personal exposure and outdoor fine particles when the relationship is not complicated by significant indoor sources.

Source: Suh *et al*. (1993). (After U.S. EPA, 1996a; Figure 7-27).

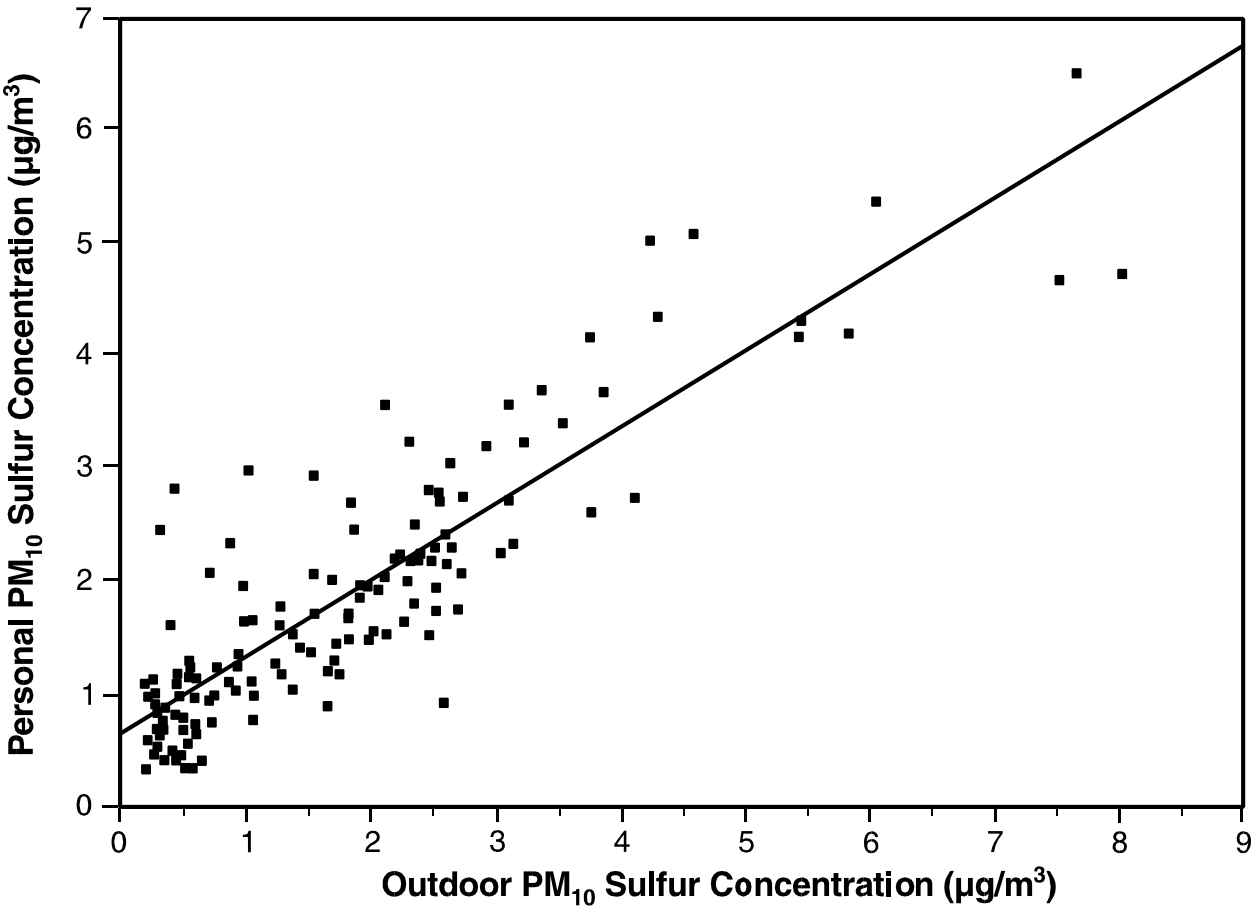


Figure D-10. Personal exposures to sulfur vs. residential outdoor concentrations: daytime $R^2 = 0.77$. Source: Ozkaynak *et al.* (1996a)

II. ANALYSIS OF POTENTIAL INFLUENCE OF MEASUREMENT ERRORS ON RELATIVE PM EFFECTS

A. Introduction

Various types of statistical models have been used to analyze the relationship of daily time series of health outcomes (such as the number of deaths or number of hospital admissions per day by cause or diagnostic code) to parallel time series of environmental predictors. The environmental predictors usually include one or more air pollutants, often specified by moving averages that may include concentrations on the response day and on one or more preceding days. The models generally also include weather variables observed on the same or preceding days and some types of adjustment for long-term temporal effects including season and other covariates.

A number of authors and commenters have recently pointed out that exposure and instrument measurement errors in environmental predictors may bias the estimated effects of the environmental factors in the daily time series regression studies used in evaluating the health effects of exposure to airborne PM, ozone, and other air pollutants. When only one predictor variable is measured imprecisely, the usual effect is to attenuate the estimated size of the effect of that predictor. This may be partly offset by changes in the estimated effects of other (more precisely measured) covariates. However, when two or more predictor variables are correlated and both are measured imprecisely, then the direction of the biases in effects estimates is less predictable, including attenuation of the estimate, occasional inflation of the estimate, or even rare reversal of its sign. The biases depend on the magnitude and correlation of measurement errors in the predictors that are imprecisely observed. In ordinary least squares (OLS) linear regression models, certain conditions favor large or unexpected biases in effect size estimates: (1) larger deviations of differences in the magnitude of the measurement errors of the standard deviations of the actual predictors, among the imprecisely measured predictors; (2) large negative correlations among predictors (sometimes large positive correlations), so that one predictor is a partial surrogate for another; (3) large positive or negative correlations among measurement errors.

In some cases, it is possible to obtain explicit analytical results that shed light on these questions. Following EPA staff discussions with Dr. Duncan Thomas at the PM Mortality Workshop sponsored by EPA in November, 1994, and thereafter, some calculations for regression coefficient bias in classical normal linear measurement error models were made by Dr. Thomas and by EPA staff. The results are reported in Marcus (1997) and tend to confirm that large biases (especially those that reverse the sign of the estimated effect) require large errors, and generally require large correlations in the error-prone predictors and large negative correlations among their measurement errors. Little is yet known about the magnitudes and distributions of actual measurement errors encountered in PM health effects studies.

Two kinds of measurement errors need to be considered: (1) "classical" errors, in which the error-prone or "noisy" observed predictor(s) W used in the regression model (such as PM_{10} or

temperature) are surrogates for the unobserved true or exact predictor(s) X; (2) "Berkson" errors, in which the predictor(s) W are known exactly, without the classical type measurement error, but are only mean values for the true exposure variable(s) X for the population as a whole. A general error model may include both classical and Berkson errors. The genesis and implications of these errors may be quite different.

B. What Kinds of "Classical" Measurement Errors May Occur?

1. Sampling and Analytical Errors

Classical measurement errors are familiar in most of the observational and experimental sciences. We assume that there is a "true" value X for some variable under study and devise a method for obtaining the value of that variable under some given set of circumstances, with the understanding that the method cannot perfectly reproduce the true value of X, but can only produce a value W that approximates X. Possible sources of "error" include sampling variability, failures in sample preparation and analysis, miscalibration of instruments, and miscoding or misreporting of results (e.g., roundoff errors). Measurement errors are often encountered in air pollution health effects research. Examples include:

a. Potential Measurement Errors in Fine Particle Mass and Coarse Particle Mass in Dichotomous Samplers.

Dichotomous particulate samplers are designed to provide direct measurements of separate PM size fractions. The sampler is typically designed to provide separate measurements of PM_{10} (thoracic particles less than 10 μm aerodynamic diameter (AD)), $PM_{2.5}$ (fine PM or FP), and $PM_{10\text{-}2.5}$ (coarse fraction of thoracic particles or CF). The sampler has two stages. The first stage is intended to remove $PM_{>10}$ particles larger than 10 μm from the air before they enter the sampler. This is accomplished by using a sampler head designed to remove 50% of particles exactly 10 μm in size, that is, to maximize the amount of PM smaller than 10 μm, and to minimize the amount of PM greater than 10μm, that enters the sampler. In reality, some particles larger than 10 μm enter the sampler, and some particles smaller than 10 μm do not enter the sampler. The degree to which these errors occur depends largely on the design of the sampler head and the rate of air flow into the sampler. The second stage is intended to separate fine PM from CF. At this stage, the intake air is split into two separate air streams. One stream, which is usually about 10% of the sampler's total air flow, entrains all CF, retains about 10% of fine PM, and impacts on a filter (filter A). The other air stream, which retains about 90% of total fine PM, impacts on a separate filter (filter B). Thus, ideally, filter A contains 100% of the CF and about 10% of the fine PM that originally entered the sampler. Ideally, filter B contains no CF and about 90% of the fine PM that entered the sampler. After correction for the percentages of total air flow that impact on filter A and filter B, and for the fact that filter A contains about 10% of the total fine PM, the tared weights of filter A and filter B provide separate measurements of CF and fine PM, respectively.

However, CF includes some particles larger than 10 μm and some particles smaller than 2.5 μm; CF particles are collected on a filter and weighed, and the CF concentration is measured as the ratio of the CF mass to the estimated volume of air passing through the sampler. The particles with AD smaller than 2.5 μm (known as fine particles or FP) are also collected on a filter and weighed, but the mass of FP includes some particles larger than 2.5 μm in diameter. Thus, FP and CF do not represent absolute ranges of particle size, as for example might be produced in a laboratory study requiring "monodisperse" aerosols with a very small distribution of sizes. PM_{10} is calculated as the sum of FP and CF concentrations.

The reader may have already identified many possible sources of "measurement error." Clean filters are weighed first, then filters with particles, and the particle mass is calculated as the difference of the two. Errors in weighing the filter, with and without particles, may occur. The volume of air flowing through the sampler may fluctuate or may be inaccurately calibrated, or the sampling period may be recorded incorrectly (computerized data acquisition systems are presumably less liable to such problems). Since these may occur at both FP and CF steps, the estimated PM_{10} = FP + CF may differ from that measured directly by a collocated one-stage PM_{10} monitor. While loss of FP from filters during shipment to the analytical laboratory appears minimal, there may be some loss of CF from Teflon (TM) filters during shipment (Dzubay and Barbour, 1983; Spengler *et al*., 1986). An estimated precision of +/- 5 $\mu g/m^3$ for FP and CF has been cited for the earlier measurements in the Harvard Six-Cities Study (Spengler *et al*., 1980). Dichotomous samplers have been evaluated in other side-by-side sampling, with relative variability of +/- 16% for FP and +/- 44% for CP (Camp, 1980). Rodes *et al.* (1985) report smaller instrument error for dichot samplers, i.e., 3.6 to 11.4% for FP and 2.6 to 18.3% for CF. Loss of particulate mass during and after collection, due to evaporation of semi-volatiles in or on the particles collected on the filters or failure to remove all particle-bound water during the equilibration process may also occur.

Both additive errors (e.g., +/- 5 $\mu g/m^3$) and multiplicative errors (e.g., +/- 16% for FP) may occur. The additive errors may represent relatively fixed uncertainties, such as the limits of measurement of the filter. The multiplicative or proportional errors may represent uncertainty associated with the mass of PM, such as differences in sampling location and random differences in the amount of PM passing through each stage. Even authors who are concerned with measurement error often fail to include both additive and multiplicative errors.

Alternative methods such as British Black Smoke (BS), Coefficient of Haze (CoH), and TEOM (Tapered Element Oscillating Microbalance) (TM) have different measurement issues that may be more serious than those faced by use of TSP, PM_{10}, $PM_{2.5}$ etc. from single-stage or multi-stage samplers. EPA has given careful attention to measurement issues.

b. Systematic and Random Biases in Pollutant Concentrations

Some measurement problems appear to introduce "random" differences or multiplicative factors, with measurements both larger and smaller than the true value in roughly equal proportions. Rounding concentrations to the nearest whole number is likely to appear as a random error. Other errors represent systematic deviations that affect every measurement. One example is the measurement of sulfates, where Lipfert (1995) suggests that a so-called "filter artifact" may have artificially increased measured sulfate values by at least 5 μg/m3 on average in some early TSP studies using glass fiber filters. Suppose that the true SO_4 measurement has both an additive systematic error (denoted SO_4FA for sulfate filter artifact, which is a random variable with a positive mean value), and a multiplicative random error due to analytical and sampling variability (denoted SOME). Then the observed sulfate concentration (denoted SO_4obs) is related to the true sulfate concentration (denoted SO_4) as:

$$SO_4obs = SO_4FA + SO_4 * SOME.$$

If the SO_4 analyses are well calibrated apart from the additive error, then the median or geometric mean of SOME is 1.

This may then propagate to other analyses. For example, suppose we are interested in both the sulfate (SO_4) and non-sulfate (NSFP) components of FP. Suppose that FP has a multiplicative measurement error (denoted FPME) from the sampling and analysis errors described previously. Thus, the observed FP (denoted FPobs) may be expressed as:

$$FPobs = FP * FPME.$$

It is generally assumed that the sulfate salts in FP are roughly equivalent to ammonium bisulfate, so that the non-sulfate components of FP should be calculated as:

$$NSFP = FP - 1.2\ SO_4.$$

However, the "observed" value (denoted NSFPobs) is derived from other observed values, thus:

$$NSFPobs = FPobs - 1.2\ SO_4obs$$
$$= FP * FPME - 1.2\ SO_4FA - 1.2\ SO_4\ SO_4ME.$$

The value of NSFPobs used in regression analyses contains all of the measurement errors FPME, SO_4FA, and SO_4ME.

c. Measurement Errors in Copollutants and other Covariates

One of the more significant issues in assessing measurement error effects arises from the fact that other criteria air pollutants are often genuinely correlated with PM, and have their own measurement error structure that may or may not be correlated with PM measurement errors. The correlations between PM and copollutants arise from several causes: (1) sources that

produce PM may also produce other gaseous pollutants, including O_3, SO_2, CO, and NO_2; (2) weather conditions that are associated with high PM concentrations, such as atmospheric inversions, may also be associated with higher concentrations of certain other pollutants; (3) some gaseous pollutants are precursors of components of PM, such as SO_2 forming sulfates and NO_2 forming nitrates. It is possible to examine the correlations between PM mass, PM components and gaseous copollutants. Careful statistical analyses may allow a significant degree of separation of the effects of PM from some of the other pollutants, at least in some locations and seasons. However, little is known about the possible correlations of measurement errors of FP, CF, and gaseous pollutants. There is considerable evidence that FP has a reasonably uniform distribution on a regional scale (Burton *et al*., 1996; Wilson and Suh, 1997), whereas CF may be less uniformly distributed. One might expect that gaseous pollutants with large spatial variability, such as CO, would be more strongly correlated with CF produced by the same local sources as CO than with regional pollutants such as FP or O_3. If the wind direction is from the CO + CF source toward the monitor, then both CO and CF would be higher than the regional average, and if the wind direction is away from the source, the both CO and CF would be lower than the average. It is, therefore, conceivable that measurement errors in PM components and copollutants may be correlated, along with the pollutants themselves, if not adjusted for underlying common factors. In relation to the above example illustrated under Subsection B, if both SO_4obs and NSFPobs are used in an epidemiological study, the measurement errors of both components could be correlated because of the occurrence of SOME in both variables. If humidity is very high when FP is measured, then the factor "1.2" used to estimate NSFP may need to be adjusted by a "random" factor to account for retention of particle bound water in hydrated sulfate salts, thus introducing an apparent correlation between humidity and NSFP, and between humidity measurement errors and NSFP measurement error.

While it is frequently assumed the meteorological variables are measured without significant error, this is not necessarily the case. Most U.S. investigators have used data from the National Climatic Data Center, which are usually measured at an airport weather station located at a considerable distance from the population center. The airport may have a substantially different micro-climate than the population center. A few hypothetical examples may illustrate the case: (1) Chicago O'Hare airport, located far to the west of the population that lives closer to Lake Michigan; (2) Los Angeles International Airport, located on the Pacific Ocean at the western edge of the L.A. population and separated from significant subpopulations by mountains (e.g., San Fernando Valley). Meteorological variables may be as likely to suffer measurement errors from sampling location and temporal variability as are some air pollutants, thus (at least hypothetically) may have measurement errors correlated with PM or with some copollutants used in epidemiological studies.

d. Measurement Errors for Every-Other-Day PM Sampling

Difference in atmospheric residence times between FP and CF is not likely to play a large role in epidemiological analyses. So long as the emissions of FP and CF occur at roughly a constant rate over periods of time of several days, one would expect an approximate steady-state concentration of each, reflecting a balance between the rate of new emissions and the concurrent rate of atmospheric dispersion. In many locations, correlations between fine-mode

(accumulation mode) and CF particles are low because the distinct physical sources of the particles do not emit FP and CF at the same rate during the course of a year. While some correlation is attributable to the common effects of weather, the correlations are not high in most SMSA's because of the diversity of types of sources. The larger temporal coefficient of variation in CF over time may reflect the possibility that there really is more temporal variability in major CF-emitting sources than in the major FP-emitting sources, rather than measurement error.

One $PM_{2.5}$ and CF study (Schwartz, et al., 1996) used a 2-day averaging period because, during much of their study, PM data were only collected every other day. Given the higher likelihood of a 2-day persistence for FP compared to CF, there may be an additional source of measurement error for CF and an increased statistical bias in favor of FP (Lipfert and Wyzga, 1997). Two-day moving averages based on alternate-day PM data could assign to each day's response an estimated PM exposure pattern for {PM data today, PM data yesterday}. If '1' represents presence, and '0' absence, the four possible PM measurement patterns are {1,1} (data on both days), {1,0} (data today, not yesterday), {0,1} (no data today, data yesterday), and {0,0} (no data). Adding indicator variables for these conditions might be sufficient to see if this assignment has any effect. However, since air pollution episodes and synoptic climatic categories typically have a duration of several days, it is likely that the amount of measurement error added to FP or CF is probably small (much less than the range of concentrations in the population) and thus will likely have minimal effects on risk estimates for FP. The situation is not as clear for CF.

2. Model Specification Errors

Mortality or hospital admissions rate models in most daily time series studies are based on small numbers of events that have a Poisson or hyper-Poisson distribution with a log-link linear model. These are fitted by likelihood or quasi-likelihood methods such as Generalized Estimating Equations (GEE). Preliminary fitting of time trends and weather effects by parametric, semi-parametric, or non-parametric methods such as smoothing splines are often used to produce "zero-mean residuals" for use in evaluating short-term PM and copollutant effects.

Misspecification of models may also play an important but unknown role when pollutant measurement errors are present. Nonlinear response models such as piecewise linear functions (including U-shaped models and threshold models) are known to be distorted by measurement errors, the shape of the curve tending to be linearized or otherwise flattened. Measurement errors can be additive, multiplicative, or a combination of additive and multiplicative, as in the example for NSFP. It may be convenient to assume a specific distributional model for predictors, such as normal, lognormal, Beta or Gamma distributions, a binomial distribution for binary covariates ("misclassification"), and so on, although there is often little basis for the specified distributional model. The effects of misspecification, whether of the functional form of the covariate model, or of the distributional model assumed for response, covariates, and measurement errors, may not be equally serious for all applications (e.g., hazard identification vs. effect size estimation vs. prediction vs. compliance testing).

3. Hypothetical Effects of Classical Measurement Error in Various Applications

The effects of classical measurement are not equally serious for all applications of statistical methods used by EPA. Consider the following uses of statistical methods:

(1) Hazard identification, by finding statistically significant effects of PM or other factors in the analysis of an epidemiological study by analysis of variance, regression models, or other statistical methods in which the relationship of health effect and pollutant is adjusted for other covariates and potential confounding factors;
(2) Quantification of concentration-response relationships for known or suspected hazards;
(3) Providing a common basis for comparing or combining the results of different studies ("meta-analyses");
(4) Using concentration-response models to predict the effects of potential future exposures;
(5) Assessment of the effects of potential Agency actions by use of "inverse regression" of a statistical regression model or meta-analysis model to calculate risks.

The effects of measurement error in the first two uses may be serious. Recent monographs (Fuller, 1987; Carroll *et al*., 1995) summarize many of the known consequences. In many linear or exponential regression models, and in similar models such as logistic regression, by far the most common effect is to attenuate (bias towards zero) the estimated regression coefficient and to increase the uncertainty associated with the estimated regression coefficient for a single error-prone variable. This always occurs with OLS linear regression models, and with logistic regression models unless the response probabilities are close to 0 or 1. When several covariates are present, and more than one of them may have classical measurement errors, then a much more complicated pattern emerges. In most cases, the coefficients are attenuated. In some cases, the coefficient may actually be inflated (biased away from zero). In other cases, the sign of the coefficient may be reversed. A detailed assessment of these effects, when true predictors are correlated and when their measurement errors are correlated, has been evaluated for OLS linear regression (Marcus, 1997). It is often possible for a strong predictor with little or no measurement error to capture some of the apparent effect of an equally strong predictor with large measurement error ("large" usually means that the measurement error standard deviation is comparable in magnitude to the population standard deviation of the true predictor). Little is known in general about the Poisson or hyper-Poisson exponential regression models when there are several error-prone predictors with realistic error models.

One of the more serious potential consequences of measurement error is that it may cause improperly-specified models to give misleading results. Nonlinear response models such as piecewise linear functions (including U-shaped models and threshold models) are known to be distorted by measurement errors, typically linearized or otherwise flattened (Yoshimura, 1990; Lipfert and Wyzga, 1997; Carroll and Galindo, 1998).

Hazard identification is affected by the increased uncertainty of regression coefficients as well. The apparent statistical significance of the effects of PM, copollutants, or weather variables may be substantially decreased because, most often, the coefficient is attenuated towards zero while its estimated uncertainty or confidence interval length is increased. In a classical significance

test, the numerator (coefficient estimate) is decreased while the denominator (estimated standard error) is increased by measurement error, so that the test statistic is greatly deflated. This is one reason why EPA has also considered some studies with "marginal" statistical significance (one-tailed P between 0.025 and 0.05, two-tailed P between 0.05 and 0.10) as useful supporting evidence: If measurement error is present, then the coefficient is even more significant than indicated by the P value, in most circumstances.

Predictive uses of regression models may be less sensitive to measurement error.
A very comprehensive summary of modern approaches to statistical inference in regression models when measurement error is present in the predictors has been published by Carroll *et al*. (1995). They conclude:

> If a predictor X is measured with error, and one wants to predict a response *based on the error-prone version W of X*, then it rarely makes sense to worry about measurement error. The one situation requiring that we model the [distribution of] measurement error occurs when we develop a prediction model using data from one population, but we wish to predict in another population. A naive prediction model that ignores measurement error may not be transportable.

As noted above, the complete characterization of exposure measurement error includes the measurement error structure of important covariates or modifying factors, such as weather. Since this is likely to vary significantly from place to place, predictive models for one SMSA that include weather-related factors may not be transportable to some other SMSA. However, to the extent that coefficients for PM_{10} and $PM_{2.5}$ appear to be more similar to each other for different SMSA's than do coefficients for weather variables, the estimated effects for PM factors may be more transportable.

C. What Kinds of "Berkson" Errors May Occur?

1. Berkson Errors and Differential Measurement Error

One of the foundations of the regulatory assessment of health effects and ecological effects from air pollution is the assumption that one or a few stationary air monitors (SAM) are sufficient to characterize exposure for most subjects and most environmental receptors in an urban or regional airshed. It has long been recognized that this introduces uncertainty in the characterization of individual exposures, but it has generally been assumed that the exposure measurement uncertainties could be ignored at the population level. The following arguments have been advanced: (1) exposure measurement uncertainty is not likely to cause a serious systematic bias in the estimated relationships of human health effects to air pollution, since exposure for most human populations tends to be averaged out by movement of individuals within the region being studied; and (2) based on regression models for normally distributed measurements, the typical effect would be an attenuation of the estimated relationship between the health effect and the air pollutant. If there were any effect of exposure measurement error, it would most likely be to *underestimate* the strength of the relationship, so that empirical relationships that were not adjusted for measurement error would at worst understate the health effects of air pollution.

The recent vigorous scientific discussions concerning the proposed revisions in NAAQS for PM and ozone have suggested that these assumptions need to be further evaluated. In particular, the following counter-arguments have been proposed: (1) differential measurement error in different locations within the region may indeed cause systematic biases in regression models, and these biases are likely to be amplified because of different patterns of air pollution exposures for various subpopulations. In particular, residential location, immigration or length of time at residence, socioeconomic status, race or ethnicity, occupation, age, and pre-existing health status may be associated with differential patterns of air pollution exposure; (2) both the true covariates and their errors of measurement may be correlated in a regression model, and certain patterns of correlation may modify the apparent relationship of health effect and air pollutant, including inflating the estimated effect or even reversing its sign, as well as attenuating the estimated effect. This is likely to be important because some pollutants are known to be highly correlated (e.g., PM and SO_2; PM and ozone in summer, in some locations); (3) personal exposure monitors or a network of residential indoor and outdoor monitors would provide more appropriate individual exposure indicators than use of one or a few SAM in each region.

Many of the commenters who have advanced these counter-arguments have failed to note the exposure measurement errors being considered here have a rather different kind of structure than so-called "classical" measurement errors. These exposure measurement errors are often called Berkson errors, that is, assigning a single value (such as the concentration at a single SAM or the average of several regional SAM concentrations) to the entire population. (See Berkson, 1950; Fuller *et al*., 1987; Carroll *et al*., 1995). Under the standard normal linear regression model, the Berkson errors do not bias the statistical inferences that can be made based on the non-adjusted classical least squares model, *provided that one can assume that the population average exposure to pollution from outdoor sources is equal to the ambient SAM level*. Reasonable

modifications allow the theory to be extended to moderately nonlinear models with heteroscedasticity, such as models with quadratic variance functions. The lack of significant bias also extends to multiple pollutant models, whether or not their Berkson measurement errors are correlated.

However, there may well be situations in which population mean exposure to outdoor pollution and ambient levels at SAM are not equivalent. There may be systematic differences due to the SAM monitoring site(s) being unrepresentative of population exposure levels. Systematic differences that are constant over time will introduce proportional biases in air pollution regression coefficients, but will not affect the apparent statistical significance of the relationship. If the relationship of air pollution to health varies randomly in a manner that is not correlated with the SAM measurements, then the Berkson error produces an additional source of variability similar to instrument errors and other analytical measurement errors. The possibility that needs to be evaluated is that of an association between ambient PM level and exposure. One hypothetical possibility, for example, is that in which exposure to CF (= PM_{10} - $PM_{2.5}$) from a point or line source varies with wind direction, which is correlated with other meteorological variables that may affect exposure-related behavior such as travel patterns or time spent outdoors.

Data are needed for characterizing uncertainty in population exposure by use of one or a few SAM in an SMSA or urban/regional airshed, with particular reference to characterizing communities in the U.S. The question is that of estimating the extent to which non-differential exposure error may be associated with spatial or temporal variations in PM exposure, with particular attention to the different PM size components such as PM_{10}, $PM_{2.5}$, $PM_{10-2.5}$, etc. It seems likely that there is not sufficient data to completely characterize these differences across any SMSA, especially when combined with variations and intra-regional differences in meteorology, in time spent by population subgroups in various micro-environments associated with different PM exposures, and especially in differences associated with susceptible sub-populations such as the elderly, or those with pre-existing COPD. Estimation of Berkson errors associated with monitor location are of particular interest.

2. Other Covariates May Also Introduce Berkson Errors

One can use ambient temperature as an example of an "ecological" variable that has commonly been used as a covariate or predictor in multivariate statistical models for daily mortality, hospital admissions, and other short-term effects. Temperature, humidity, and other meteorological variables are measured at a location that is not necessarily related to population "exposure", for example, at a metropolitan airport that may be many miles away from the center of the residential or work sites of the SMSA population represented by these variables. Local micro-environmental conditions may differ substantially from those at the airport weather station, including mitigating effects on temperature, humidity, wind speed, etc. related to proximity to bodies of water, to parkland or forest, to hills and valleys, to urban center heat islands, to urban street canyon effects, and so on.

Individual exposure to "ambient temperature" may be even less adequately characterized by a single SMSA measurement than is individual exposure to some air pollutants. For example, most individuals spend most of their time indoors. The preferred range of temperatures for most people is likely to be in a very narrow range, say 20 to 24 degrees C. (68 to 75 degrees F.), which can be substantially regulated in almost all indoor locations by adjustments of heating and air conditioning, throughout all seasons. This is surely a smaller range of indoor temperature conditions than the likely range of indoor air pollution concentrations for some components, particularly fine particles, that readily penetrate into the indoor environment and are not readily removed in that environment. Outdoor exposures to ambient temperature can also be regulated by behavioral adjustments. These adjustments include selections of clothing appropriate to the weather, use of private or public transportation in which temperature "exposure" can be regulated as it is indoors by heating and air conditioning, and by averting behavior in which exposure to temperature extremes may simply be avoided by delaying outdoor activities. No such averting behavior is likely when ambient air pollution concentrations are "high" by current exposure levels. Therefore, the ambient regional pollution concentrations may be more representative of some air pollution exposures, such as $PM_{2.5}$, than are ambient regional temperatures likely to be representative of individual exposures to ambient temperature. While the argument that regional ambient temperature measurements cannot be used to characterize individual temperature exposure may be somewhat hypothetical, it is at least as plausible as the commenter's argument that ambient regional measurements of $PM_{2.5}$, PM_{10}, and other air pollutants cannot be used to characterize individual air pollution exposures.

EPA has not considered measurement error in personal exposure to temperature or other weather variables in the PM CD because these considerations are expected to have little or no impact on the use of the PM epidemiological models for risk assessment and risk management applications.

3. Personal Exposure Monitoring and Berkson Errors

The Berkson model arises from the fact that the investigator in doing daily time series analyses assigns the same PM value to characterize the personal exposure concentration of ambient origin for all individuals in the population, whereas the actual personal exposure to ambient PM (which we denote Ep,a, see equation 5 above) is somewhat different for each of the individuals. For air pollutants with a regionally uniform distribution, such as $PM_{2.5}$ or sulfates, and PM_{10} to a lesser extent, for which PM is a reasonably unbiased estimate of average population ambient exposure, the effect on statistical estimation may be minimal. For the statistical reader, we would describe the additive Berkson model as true Ep,a = PM + error, E(Ep,a | PM) = PM, where E( ) is the average over the exposed population. Then, as Carroll *et al*. (1995, p. 150) point out: "The Berkson additive model has an unusual feature, in that for linear regression the naive analysis ignoring measurement error gives correct inference about the regression ..." It is likely that this condition applies to measured ambient PM as an unbiased estimate of Ep,a, which does not require that all exposed individuals have exactly the same ambient personal exposure concentration to PM of ambient origin.

As discussed in section I of this appendix, one must be careful to distinguish the personal exposure to ambient PM from the total personal PM exposure. As shown in Figure D-2 above,

the total personal exposure includes indoor exposures, both from particles that are generated indoors and from particles generated outside the residence that infiltrate the residence. The personal exposure to ambient PM includes exposure to ambient particles outdoors and exposure indoors to ambient particles that have penetrated indoors. While the concentration outside the residence (which we denote Ca) may differ from the personal exposure to ambient PM, it is probable that the regional PM indicators being evaluated are also reasonably unbiased estimators of Ca, and thus proportional to the fraction of Ca that penetrates the residence. As discussed above , there is substantial evidence that concentrations of particles generated by indoor sources are almost completely independent of outdoor (ambient) particle concentrations, and are therefore not likely to be a confounder in health studies of exposure to particles generated outdoors.

The other measurement error is the difference between the measured and the true ambient concentration at the central air monitoring station(s), which denoted as AC. In this case, the classical measurement error model (PM = AC + error) may be applicable if marked concentration gradients exist in the community, and further attention is required to quantify the effects of such errors. EPA's assessment is that "ambient concentration measurement errors" are unlikely to have such large magnitude in a non-industrial urban area or such strong correlations with weather or with co-pollutants are unlikely to induce substantial quantitative modifications of effects size estimates associated with exposure to ambient PM.

D. Measurement Errors in PM Health Effects Studies

1. Potential Effects of Measurement Errors in PM Studies

Lipfert and Wyzga, in particular, have raised the possibility that measurement errors may distort observed exposure-response relationships, in that such errors may make a non-linear (threshold) type underlying relationship appear to be linear (no thresholds) and/or that relatively stronger contributions to observed associations with health endpoints may be misattributed to PM versus less precisely measured copollutants or to more precisely measured size-specific PM fractions verses less precised measured PM components.

Distortion of the concentration-response curve is a theoretical possibility. If the response function is strongly nonlinear, then some distortion may occur for exposure measurement errors of Berkson type as well as the classical errors-in-variables model. However, there is currently little evidence that the response function has more than modest nonlinearity at current levels of PM and other pollutants below the current NAAQS.

Strong confounding by other copolllutants or weather cannot occur unless the confounder has at least as strong an effect on response as does PM, is highly correlated with PM, and the measurement error in these variables is comparable to the population range. Strong correlation means $r > 0.9$ or $r < -0.9$; such high correlations are rarely observed. HEI studies (Samet *et al*., 1996b) suggest that weather effects can be largely separated from those of various air pollutants, and that some pollutants can be separated from others. Ozone is a seasonal pollutant that may

have mortality effects, but these effects are often separable from other pollutants; lagged CO may also have a large independent effect that is somewhat separable from other pollutants. The Philadelphia data showed that the effects of TSP and SO_2 were harder to separate statistically, although there is little physiological reason to believe that SO_2 acting directly causes rapid mortality. NO_x appeared to be part of the TSP-SO_2 group, but could be dismissed because NOx produced much weaker effects by itself than did either TSP or SO_2, and in a combined model, the NO_x effects were estimated as weakly beneficial (which is obviously a false finding). There is little reason to believe that the whole of the health effects attributed to PM in time series analyses are quantitatively explained by measurement errors in confounding variables.

2. The Role of Indoor-Generated PM as an "Exposure Error" in PM Epidemiological Studies

Personal exposures to PM generated by indoor sources are largely independent of exposures to ambient PM from outdoor sources, and thus largely independent of the component of indoor personal exposure attributable to infiltration of ambient PM. Since the NAAQS are intended to protect individuals against the component of air pollution exposures from outdoor-generated PM and other criteria pollutants, it is only the outdoor-generated component of personal exposures to FP and CP that is relevant to EPA assessment of ambient PM health effects.

The real issue is whether exposure measurement error seriously distorts the attribution of health effects additively (that is linearly) across indoor- and outdoor-generated PM. This is possible if the concentration-response function is strongly nonlinear, a threshold function for example, with the threshold somewhere in the range of PM levels being considered for the proposed standards. In that case, one could hypothesize that the effects of ambient PM might be very different, depending on the level of indoor-generated PM. For example, suppose that the population is essentially bimodal, one subpopulation with a very high level of indoor-generated PM exposure (from smoking cigarettes, for example), and another subpopulation with a very low level of indoor-generated PM (nonsmokers, for example). Then it is possible that the first subpopulation may be pushed above the (purely hypothetical) response threshold by additional exposure from ambient PM in addition to the indoor-generated personal exposure, whereas the second subpopulation is exposed only to ambient PM from outdoor sources, and the total personal exposure is below the (purely hypothetical) response threshold. Exposure misclassification may distort the true exposure-response relationship. However, no convincing evidence has been presented to EPA which establishes the likelihood of this.

In summary, indoor-generated PM may cause additional adverse health effects, but these effects are likely independent of and may often be additive to effects from ambient PM. Time-series epidemiological studies using changes in ambient PM and other environmental stressors as predictors of changes in health response are not likely to be seriously biased by the indoor-generated component of personal exposure.

3. Possible Effects of Berkson Errors in Weather Variables

Temperature errors are a candidate for Berkson-structure measurement errors that may need to be considered, if measurement errors in pollution are being evaluated, because of the virtual certainty that the response of population mortality rates to temperature is a strongly non-linear function. Temperature exposure error arises when the measured temperature value at the regional weather station is assigned as the daily ambient exposure to temperature of the whole population. In population-based studies, where the response is the hospital admissions count or mortality count in the population, this may be an appropriate example of the kind of measurement error called a ***Berkson error model***. This is discussed in basic monographs such as *Measurement Error Models* by W. Fuller (J. Wiley, N.Y., 1987, pp. 82-83), and in *Nonlinear Measurement Error Models* by R. Carroll, D. Ruppert, and L. Stefanski (Chapman and Hall, N.Y. and London, 1995, pp. 3, 9, 52-56, 67, 143, 150-151, 255), with extensive references to earlier papers. The Berkson model arises from the fact that the investigator doing daily time series analyses assigns the same temperature value to characterize the ambient personal exposure concentration for all individuals in the population, whereas the actual Ambient Personal Temperature Exposure (which we denote APTE) is somewhat different for each of the individuals. If exposure to ambient temperature has a regionally uniform distribution, the effect on covariate adjustment for statistical estimation of weather effects, and their influence on PM effects, may be minimal. As Carroll *et al*. (1995, p. 150) point out: "The Berkson additive model has an unusual feature, in that for linear regression the naive analysis ignoring measurement error gives correct inference about the regression ..." It is likely that this condition applies to measured ambient temperature as an unbiased estimate of APTE, which does not require that all exposed individuals have exactly the same ambient personal temperature exposure.

While recognizing that more measurement error research would be useful, EPA's position is that no adjustment of temperature and other weather variables is needed for the Berkson-type measurement error structure in daily time series analyses. Daily time series studies relate day-to-day changes in response (such as mortality) to changes in air pollution and weather over the preceding few days. Whatever variables are used to characterize the weather, responses to these changes (which occur during a small part of any given season) are not likely to be strongly non-linear, even though strongly non-linear functions for response may characterize the response function over the entire year. The short-term nature of the time series concentration-response studies legitimizes the assumption of Berkson measurement error structure for the population, which requires no change in statistical methodology from what is commonly used. While analytical and instrumental measurement errors may occur for weather variables, they are almost certainly of little quantitative significance in regression models. It is likely that any of several model specifications for weather effects can be used for time series analyses, with little potential for confounding of air pollution effects, and little potential for classical or Berkson measurement error effects.

E. Conclusions

In summary, several commenters have pointed out that, in multiple pollutant analyses, measurement error or, more generally, exposure misclassification, could theoretically bias effects estimates of PM or co-pollutants in either direction, introducing further uncertainties into the estimated concentration-response relationships for all pollutants. EPA believes that the quantitative importance of many of these uncertainties have been exaggerated. While there may be substantial inter-personal differences in exposure to ambient PM, there is little reason to believe that measurements made at one or several community air monitoring stations and at a single weather station do not adequately characterize day-to-day changes in average population exposure to ambient air pollutants and weather. Therefore, time series analyses relating regional health response to regional air pollution and weather are expected to show little response to variability in personal exposure.

Furthermore, as determined by EPA's expanded statistical analyses noted above, three conditions must exist before measurement error can generate serious bias of the types hypothesized by the commenters. Specifically, these are:

(1) The measurement error in the poorly measured pollutant must be very large, roughly at least the same size as the population variability in that pollutant;

(2) The poorly measured pollutant must be highly correlated with the other pollutant, either positively or negatively; and,

(3) The measurement errors for the two pollutants must be highly negatively correlated.

Commenters have not provided evidence to substantiate that such conditions actually occur with respect to the measurement of ambient PM in relation to those for gaseous co-pollutants commonly used in epidemiological studies. Therefore, it appears unlikely that measurement and exposure errors for PM and other pollutants have inflated the estimated effects of PM, even in multivariate analyses.

More importantly, the available evidence on the consistency of the PM-effects relationships in multiple urban locations, with widely varying indoor/outdoor conditions and a variety of monitoring approaches, makes it unlikely that the observed associations of PM with serious health effects at levels allowed under the current NAAQS are an artifact of errors in measurement of ambient pollution or of human exposure.

REFERENCES:

Abbey, D.E.; Mills, P.K.; Petersen, F.F.; Beeson, W.L. (1991) Long-term ambient concentrations of total suspended particulates and oxidants as related to incidence of chronic disease in California Seventh-Day Adventists. Environ. Health. Perspect. 94:43-50.

Abt Associates (1996a) Proposed Methodology for PM Risk Analyses in Selected Cities (Draft). Prepared by Abt Associates, Inc., Hampden Square, Suite 500, 4800 Montgomery Lane, Bethesda, MD 20814-5341 12 February 1996.

Abt Associates (1996b) A Particulate Matter Risk Analysis for Philadelphia and Los Angeles. Prepared by Abt Associates for US EPA, OAQPS, Hampden Square, Suite 500, 4800 Montgomery Lane, Bethesda, MD 20814-5341 3 July 1996 (Updated November 1996).

Abt Associates (1997a) Revision of Mortality Incidence Estimates Based on Pope et al. (1995) in the Abt Particulate Matter Risk Assessment Report. Prepared by Abt Associates, Inc., Hampden Square, Suite 500, 4800 Montgomery Lane, Bethesda, MD 29814-5341. 5 June 1997.

Abt Associates (1997b) Revision of Mortality Incidence Estimates Based on Pope et al. (1995) in the Abt Particulate Matter Risk Assessment Report. Prepared by Abt Associates, Inc., Hampden Square, Suite 500, 4800 Montgomery Lane, Bethesda, MD 29814-5341. 6 June 1997.

Anderson HR, Spix C, Medina S, Schouten JP, Castellsague J, Rossi G, Zmirou D, Touloumi G, Wojtyniak B, Ponka A, Bacharova L, Schwartz J, Katsouyanni K. 1997. Air pollution and daily admissions for chronic obstructive pulmonary disease in 6 European cities: results from the APHEA project. Eur Respir J 10:1064-1071

Anderson HR, Ponce de Leon A, Bland JM, Bower JS, Strachan DP. 1996. Air pollution and daily mortality in London: 1987-92. BMJ 312:665-9

Bacharova L, Fandakova K, Bratinka J, Budinska M, Bachar J, Gudaba M. 1996. The association between air pollution and the daily number of deaths: findings from the Slovak Republic contribution to the APHEA project. J Epidemiol Comm Health 50(Suppl 1):S19-S21

Balmes et al. (Unpublished abstract) Attached to comments received from Pacific Gas and Electric Company, Docket No. A-95-54, IV-D-2183. March 11, 1997.

Becker, S., J.M. Soukup, M.I. Gilmour, and R.B. Devlin (1996). Stimulation of human and rat alveolar macrophages by urban air particles: effects on oxidant radical generation and cytokine production. Toxicol. Appl. Pharmacol. 141:637-648.

Bennett, W.D., Zeman, K.L., Kim, C.S. (1996). Variability of fine particle deposition in healthy adults: effects of age and gender. Am J Respir Crit Care Med 153:1641-1647.

Bennett, W.D., Zeman, K.L., Kim, C., Mascarella, J. (1997) Enhanced deposition of fine particles in COPD patients spontaneously breathing at rest. Inhal. Toxicol., 9:1-14.

Berkson, J. (1950) Are there two regressions? J. Amer. Statistical Assoc. 45:164-179.

Binkowski, F.S.; S.M. Kreidenweis,; D.Y. Harrington, U. Shankar (1996) Comparison of new particle formation mechanisms in the regional particulate model. Presented at the fifteenth annual meeting of the American Association for Aerosol Research, Orlando, FL, October 14-18, 1996.

Binkowski, F.S. and Shankar, U. (1995) The regional particulate model: Part I Model description and preliminary results. J. Geophys. Res., 100:26191-26209.

Borja-Aburto, V.H.; Loomis, D.P.; Bangdiwala, S.I.; Shy, C.M.; Rascon-Pacheco, R.A. (1997) Ozone, suspended particulates, and daily mortality in Mexico City. Am. J. Epidemiol. 145:258-268.

Braun-Fahrlander, C.; Ackermann-Liebrich, U.; Schwartz, J.; Gnehm, H.P.; Rutishauser, M.; Wanner, H.U. (1992) Air pollution and respiratory symptoms in preschool children. Am. Rev. Respir. Dis. 145:42-47.

Brown, K. (1997) Predictability of personal exposure to airborne particulate matter from ambient concentrations in four communities. In: Comments of the American Petroleum Institute on the proposed national ambient air quality standards for particulate matter (Appendix F). Docket No. A-95-54, IV-D-2247.

Burnett, R.T., Dales, R.E., Brook, J.R., Raizenne, M.E., Krewski, D. (1997) Association between ambient carbon monoxide levels and hospitalizations for congestive heart failure in the elderly in 10 Canadian cities. Epidemiol 8:162-167

Burnett, R.T.; Dales, R.E.; Raizenne, M.E.; Krewski, D.; Summers, P.W.; Roberts, G.R.; Raad-Young, M.; Dann, T.; Brook, J. (1994) Effects of low ambient levels of ozone and sulfates on the frequency of respiratory admissions to Ontario hospitals. Environ Res. 65:172-194.

Burnett, R.T.; Merange, J.C.; Franklin, C.A. (1989) Environ Res 49(1):20-39.

Burton, R.M., Suh, HH., Koutrakis (1995). Spatial variation in particulate concentrations within metropolitan Philadelphia. Environ Sci Technol 30:400-407

Carroll, R.J., Ruppert, D., Stefanski, L. (1995) Measurement Error in Nonlinear Models. Chapman and Hall: London

Carroll, R.J., Galindo, P. (1996) Effects of measurement error on the validation of complex computer models for childhood lead exposure. Paper presented at Lead Model Validation

Workshop sponsored by USEPA in Chapel Hill, NC, Oct. 21-23, 1996. Accepted for publication: Environmental Health Perspectives

Carter, J.D., A.Ghio, J.Samet, and R.B. Devlin. (1997) Cytokine production by human airway epithelial cells after exposure to an air pollution particle is metal dependent. Toxicol. Appl. Pharmacol. (In Press).

Chestnut, L.G.; Dennis, R.L.; Latimer, D.A. (1994) Economic benefits of improvements in visibility: acid rain provisions of the 1990 Clean Air Act Amendments. Proceedings of the aerosols and atmospheric optics: radiative balance and visual air quality. Air & Waste Management International Specialty Conference, pp. 791-802.

Chestnut, L.G.; Rowe, R.D. (1990) Preservation values for visibility in the national parks. Washington, DC: U.S. Environmental Protection Agency.

Choudbury AH, Gordian ME, Morris SS. (1997) Associations between respiratory illness and PM10 air pollution. Arch Environ Health 52:113-117

Churg, A. and Brauer, M. (1997) Human lung parenchyma retains PM2.5. Am J Respir Crit Care Med 155:2109-2111.

Cifuentes, L.; Lave, L.B. (1996) Association of daily mortality and air pollution in Philadelphia, 1983-1988. Doctoral dissertation: Carnegie-Mellon University

Cooper, J.A.; Tawney, C.W. (1996) Comments on OAQPS recommended PM2.5 concentrations equivalent to current PM10 standards. In: Comments of the Utility Air Regulatory Group (Attachment 6). Docket No. A-95-54, IV-D-2250. March 12, 1997.

Costa, D.L.; Dreher, K.L. (1997) Bioavailable transition metals in particulate matter mediate cardiopulmonary injury in healthy and compromised animal models. Environ. Health Perspect. (In press)

Crandall RW, Rueter FH, Steger, WA (1996) Clearing the Air: EPA's Self-assessment of Clean-Air Policy. Regulation, Number 4, 35-46.

Dab W, Medina S, Quenel P, Le Moullec Y, Le Tertre A, Thelot B, Monteil C, Lameloise P, Pirard P, Mornas I, Ferry R, Festy B. 1996. Short term respiratory health effects of ambient air pollution: results of the APHEA project in Paris. J Epidemiol Comm Health 50(Suppl 1):S42-S46

Damberg, R.; Polkowsky, B. (1996) Methodology for Estimating Visibility Improvements Due to Reductions in Fine Particle Concentrations. Memorandum to the docket for the PM NAAQS review. September 1996.

Davis, J.M.; Sacks, J.; Saltzman, N.; Smith, R.L.; Styer, P. (1996) Airborne particulate matter and daily mortality in Birmingham, Alabama. Technical Report #55, National Institute of Statistical Sciences, P.O. Box 14162, Research Triangle Park, NC 27709.

Delfino RJ, Murphy-Moulton AM, Burnett RT, Brook JR, Becklake MR. 1997. Effects of air pollution on emergency room visits for respiratory illnesses in Montreal, Quebec. Am J Respir Crit Care Med 155:568-576

Dockery, D.W.; Pope, C.A., III; Wu, W.; Spengler, J.D.; Ware, J.H.; Fay, M.E.; Ferris, B.G., Jr.; Speizer, F.E. (1993) An association between air pollution and mortality in six U.S. cities. N. Engl. J. Med. 329:1753-1759.

Dockery, D.W.; Cunningham, J.; Damokosh, A.I.; Neas, L.M.; Spengler, J.D.; Koutrakis, P.; Ware, J.H.; Raizenne, M.; Speizer, F.E. (1996) Health effects of acid aerosols on North American children: respiratory symptoms. Environ. Health. Perspect. 104:500-505.

Dockery, D.W.; Schwartz, J.; Spengler, J.D. (1992) Air pollution and daily mortality: associations with particulates and acid aerosols. Environ. Res. 59:362-373.

Dockery, D.W.; Speizer, F.E.; Stram, D.O.; Ware, J.H.; Spengler, J.D.; Ferris, B.G., Jr. (1989) Effects of inhalable particles on respiratory health of children. Am. Rev. Resp. Dis. 139:587-594

Donaldson, K., Beswick, P.H., Gilmour, P.S. (1996) Free radical activity associated with the surface of particles: a unifying factor in determining biological activity? Toxicology Letters, 88: 293-298.

Dreher, K.L., Jaskot, R.H., Lehmann, J.R., Richards, J.H., McGee, J.K. (1997) Soluble transition metals mediate residual oil fly ash induced acute lung injury. J. Toxicol. Environ. Health, 50: 285-305.

Dye, J.A., Adler, K.B., Richards, J.H. and Dreher, K.L. (1997) Epithelial injury induced by residual oil fly ash: role of reactive oxygen species. Am. J. Respir. Cell Molec. Biol., (in press).

Dzubay, T.G.; Barbour, R.K. (1983) A method to improve the adhesion of aerosol particles on Teflon filters. J Air Pollut Control Assoc 33:962-965

Fairley, D. (1990) The relationship of daily mortality to suspended particulates in Santa Clara County, 1980-86. Environ. Health Perspect. 89:159-168.

Fitz-Simons, T.; Mintz, D.; Wayland, M. (1996) Proposed methodology for predicting $PM_{2.5}$ from PM_{10} values to assess the impact of alternative forms and levels of the PM NAAQS. Document transmitted to members of the Clean Air Act Scientific Advisory Committee on June 26, 1996.

Fitz-Simons, T. (1997) Memorandum to the files: Response to comments made by AISI on EPA methodology for predicting $PM_{2.5}$ from PM_{10}. June 5, 1997.

Freas, W. (1997) Memorandum to the files: Ratios of particulate matter annual arithmetic mean to annual geometric mean concentrations. Docket No. A-95-54, IV-B-2. May 29, 1997.

Frew, A. J.; Salvi, S. S. (1997) Diesel exhaust particles and respiratory allergy [editorial]. Clin. Exp. Allergy 27: 237-239.

Fuller, W. (1987) Measurement Error Models. John Wiley, NY.

Gallant A.R., Goebel, J. J. (1976) Nonlinear regression with autocorrelated errors. J. Amer. Statistical Assoc. 71:961-967.

Gamble, J.F.; Lewis, R.J. (1996) Health and respirable particulate (PM_{10}) air pollution: a causal or statistical association? Environ. Health. Perspect. 104:838-850.

Gavett, S.H., Madison, S.L., Dreher, K.L., Winsett, D.W., McGee, J.K., Costa, D.L. (1997). Metal and sulfate composition of residual oil fly ash determines airway hyperreactivity and lung injury in rats. Environ. Res., 72: 162-172.

GCVTC (1996) Recommendations for improving Western vistas. Report of the Grand Canyon Visibility Transport Commission (GCVTC) to the U.S. EPA. June 1996.

Ghio, A.J., Stonehuerner, Pritchard, R.J., Piantadosi, C. A., Quigley, D.R., Dreher, K.L.., Costa, D. L. (1996) Humic-like substances in air pollution particulates correlate with concentrations of transition metals and oxidant generation. Inhal. Toxicol., 8: 479-494.

Gilmour, P.S., Brown, D.M., Lindsay, T.G., Beswick, P.H., MacNee, W., Donaldson, K. (1996) Adverse health effects of PM10 particles: involvement of iron in generation of hydroxyl radical. Occup Environ Med, 53: 817-822.

Ghio, A.J., Meng, Z.H., Hatch, G.E., Costa, D.L. (1997) Luminol-enhanced chemiluminescence after in vitro exposures of rat alveolar macrophages to oil fly ash is metal dependent. Inhal. Toxicol., 9: 255-271.

Godleski, J.J., Sioutas, C., Katler, M., Catalano P., Koutrakis, P. (1996) Death from inhalation of concentrated ambient air particles in animal models of pulmonary disease. In: Lee, J., Phalen R. (eds). (Proceedings of the second annual colloquium on particulate pollution and human health effects). 4:136-143, 1996

Gordian, M.E.; Ozkaynak, H,; Xue, J.; Morris, S.S.; Spengler, J.D. (1996) Particulate air pollution and respiratory disease in Anchorage, Alaska. Environ. Health Perspect. 104:290-297.

Greenland, S. (1991) Summarization, smoothing, and inference in epidemiologic analysis. Scand J Soc Med (Ser A) 21(4):227-232.

Hatch, G.E., Boykin, E., Graham, J.A., Lewtas, J., Pott, F., Loud, D., Mumford, J.L. (1985) Inhalable particles and pulmonary host defense: in vivo and in vitro effects of ambient air and combustion particles. Environ. Res., 36: 67-80.

Hayes SR. (1991). Use of an indoor air quality model (IAQM) to estimate indoor ozone levels. Journal of the Air & Waste Management Association. 41(2):161-70.

Hefflin, B.J.; Jalaludin, B.; McClure, E.; Cobb, N.; Johnson, C.A.; Jecha, L.; Etzel, R.A. (1994) Surveillance for dust storms and respiratory diseases in Washington State, 1991. Arch. Environ. Health 49:170-174.

HEI. (1995) Particulate Air Pollution and Daily Mortality: Replication and Validation of Selected Studies (The Phase I.A Report of the Particle Epidemiology Evaluation Project). Health Effects Institute, Cambridge, MA.

HEI. (1997) Particulate Air Pollution and Daily Mortality: Analyses of the Effects of Weather and Multiple Air Pollutants (The Phase I.B Report of the Particle Epidemiology Evaluation Project). Health Effects Institute, Cambridge, MA.

Hill, A.B. (1965) The environment and disease: associations or causation? Proc. R. Soc. Med. 58:295-300.

Hill, I.R. (1996) Reactions to particles in smoke. Toxicol. 115: 119-122.

Hoek, G.; Brunekreef, B. (1994) Effects of low-level winter air pollution concentrations on respiratory health of Dutch children. Environ Res. 64:136-150.

Ito, K.; Kinney, P.L.; Thurston, G.D. (1995) Variations in PM-10 concentrations within two metropolitan areas and their implications for health effects analyses. Inhal. Toxicol. 7:735-745.

Ito, K.; Thurston, G.D.; Hayes, C.; Lippmann, M. (1993) Associations of London, England, daily mortality with particulate matter, sulfur dioxide, and acid aerosol pollution. Arch Environ Health 48:213-220.

Katsouyanni K, Touloumi G, Spix S, Schwartz J, Balducci F, Medina S, Rossi G, Wojtyniak B, Sunyer J, Bacharova L, Schouten JP, Ponka A, Anderson HR. 1997. Short-term effects of ambient sulphur dioxide and particulate matter on mortality in 12 European cities: results from the APHEA project. BMJ:(in press)

Katsouyanni K, Schwartz J, Spix C, Touloumi G, Zmirou D, Zanobetti A, Wojtyniak B, Vonk JM, Tobias A, Ponka A, Medina S, Bacharova L, Anderson HR. 1996. Short term effects of air

pollution on health: a European approach using epidemiologic time series data: the APHEA protocol. J Epidemiol Comm Health 50(Suppl 1):S12-S18

Katsouyanni, K.; Hatzakis, A.; Kalandidi, A.; Trichopoulos, D. (1990b) Short-term effects of atmospheric pollution on mortality in Athens. Arch Hellen Med. 7:126-132.

Kaufman, J.W., Scherer, P.W., Yang, C.G. (1996) Predicted combustion product deposition in the human airway. Toxicol., 115: 123-128.

Keeler, G.J.; Spengler, J.D.; Castillo, R.A. (1991). Acid aerosol measurements at the suburban Connecticut site. Atmos. Environ. Pat A 25: 681-690.

Killingsworth, C.R., Alessandrini, F., Krishna Murphy, G.G., Catalano, P.J., Paulauskis, J.D., Godleski, J.J. (1997) Inflammation, chemokine expression, and death in monocortaline-treated rats following fuel oil fly ash inhalation. Inhal. Toxicol.: (accepted).

Kim C.S., Hu, S.C., DeWitt, P., Gerrity, T.R. (1996) Assessment of regional deposition of inhaled particles in human lungs by serial bolus delivery method. J. Appl. Physiol. 81: 2203-2213.

Kim, C.S., Kang, T.C. (1997) Comparative measurments of lung deposition of inhaled fine particles in normal subjects and patients with obstructive airway disease. Am. J. Respir. Crit. Care Med. 155: 899-905.

Kinney, P.L.; Ito, K.; Thurston, G.D. (1995) A sensitivity analysis of mortality/PM10 associations in Los Angeles. Inhal Toxicol. 7:59-69.

Kinney, P.L.; Ozkaynak, H. (1991) Associations of daily mortality and air pollution in Los Angeles County. Environ Res. 54:99-120.

Kleinman, M.T.; Bhalla, D.K.; Mautz, W.J.; Phalen, R.F. (1995) Cellular and immunologic injury with PM10 inhalation. Inhal Toxicol. 7:589-602.

Knox, R.B., Suphioglu, C., Taylor, P., Desai, R., Watson, H.C., Peng, J.L. & Bursill, L.A. (1997) Major grass pollen allergen Lol p 1 binds to diesel exhaust particles: implications for asthma and air pollution. Clin. Exper. Allergy, 27: 246-251.

Kodavanti, U.P., Jaskot, R.H., Wei, Y.S., Costa, D.L., Ghio, A.J. & Dreher, K.L. (1997) Genetic variability in combustion particle-induced chronic lung injury. Am. J. Physiol. 272 (Lung Cell. Mol. Physiol. 16): L521-L532.

Kodavanti, U. P.; Jaskot, R. H.; Costa, D. L.; Dreher, K. L. (1997) Pulmonary proinflammatory gene induction following acute exposure to residual oil fly ash: roles of particle associated metals. Inhal. Toxicol.: accepted.

Koman, P.D. (1997) Memorandum to the files: Supplemental information regarding $PM_{2.5}$ air quality values from key short-term fine particle exposure studies. June 30, 1997.

Lay J.C., Bennett, W.D., Kim, C.S., Devlin, R.B., and Bromberg, P.A. Retention and intracellular distribution of instilled iron oxide particles in human alveolar macrophages. Am J Respir Cell Molec Biol. (Submitted, 1997).

Lee, H.S.; Wadden, R.A.; Scheff, P.A. (1993). Measurement and evaluation of acid air pollutants in Chicago using an annular deluder system. Atmos. Environ. Part A 27: 543-553.

Liang, K.-Y., Zeger, S.L. (1986) Longitudinal data analysis using generalized linear models. Biometrika 73:13-22

Li, Y.; Roth, H.D. (1995) Daily mortality analysis by using different regression models in Philadelphia County, 1973-1990. Inhal. Toxicol. 7:45-58.

Linn WS, Shamoo DA, Anderson KR, Peng R-C, Avol EL, Hackney JD, Gong H. 1996. Short-term pollution exposures and responses in Los Angeles area schoolchildren. J Expos Anal Environ Epidemiol 6:449-472

Lioy, PJ, Waldman, JM, Buckley, T, Butler, J, Pietarinen, C, (1990). The personal, indoor and outdoor concentrations of PM-10 measured in an industrial community in the winter. Atmos. Environ Part B 24: 57-66

Lipfert, FW, and Wyzga, RE (1997) Air Pollution and Mortality: the Implications of Uncertainties in Regression Modeling and Exposure Measurement. J Air Waste Manage Assoc 47 (4) 517-523.

Lipfert, F.W. (1997) Comments on the proposed changes to the NAAQS for ozone and particulate matter. Docket No. A-95-54, IV-D-2180. March 11, 1997.

Lipfert, F.W.; Wyzga, R.E. (1995) Air pollution and mortality: issues and uncertainties. J. Air Waste Manage. Assoc. 45:949-966.

Lipfert, F.W. (1995) Estimating air pollution-mortality risks from cross-sectional studies: prospective vs. Ecologic study designs. In: Particulate matter: health and regulatory issues: proceedings of an international specialty conference; April; Pittsburgh, PA. Pittsburgh, PA: Air & Waste Management Association; p. 78-102. (A&WMA publication VIP-49)

Lipfert, F.W.; Urch, R.B. (1997) Personal exposures of asthmatics to air pollution in Toronto and implications for epidemiological studies. Submitted as attachment to comments on the proposed NAAQS for particulate matter by R.E. Wyzga, Docket No. A-95-54, IV-D-2329, March 11, 1997.

Lippmann, M.; Shy, C.; Stolwijk, J.; Speizer, F. (1996) Letter to Administrator Carol M. Browner. Supplement to the Closure Letter from the Clean Air Scientific Advisory Committee. March 20, 1996.

Lipsett M, Hurley S, Ostro B. 1997. Air Pollution and emergency room visits for asthma in Santa Clara County, California. Environ Health Perspect 105(2):216-222

Liu, L.-J.S.; Burton, R.; Wilson, W.E.; Koutrakis, P. (1996). Comparison of aerosol acidity in urban and semi-rural environments. Atmos. Environ. 30: 1237-1245.

Lyon, J.L.; Mori, M.; Gao, R. (1995) Is there a causal association between excess mortality and exposure to PM-10 air pollution? Inhal. Toxicol. 7:603-614.

Marcus, A. (1997) Measurement error, correlation and confounding. USEPA/NCEA Docket.

Martin, A.E.; Bradley, W.H. (1960) Mortality, fog and atmospheric pollution: an investigation during the winter of 1958-1959. Mon. Bull. Minist. Health Public Health Lab. Serv. (GB) 19:56-73.

Miguel, A.G., Cass, G.R., Weiss, J., & Glovsky, M.M. (1996) Latex allergens in tire dust and airborne particles. Env. Health Perspectives, 104 (11): 1180-1185.

Ministry of Health - London- Her Majesty's Stationary Office (1954) Reports on Public Health and Medical Subjects No. 95. Mortality and Morbidity During the London Fog of December 1952.

Moolgavkar, S.H.; Luebeck, E.G.; Anderson, E.L. (1997) Air pollution and hospital admissions for respiratory causes in Minneapolis-St. Paul and Birmingham. Epidemiol. 8:364-370

Moolgavkar, S.H.; Luebeck, E.G. (1996) Particulate air pollution and mortality: a critical review of the evidence. Epidemiol. 7:420-428.

Moolgavkar, S.H.; Luebeck, E.G.; Hall, T.A.; Anderson, E.L. (1995a) Air pollution and daily mortality in Philadelphia. Epidemiol. 6:476-484.

Moolgavkar, S.H.; Luebeck, E.G.; Hall, T.A.; Anderson, E.L. (1995b) Particulate air pollution, sulfur dioxide, and daily mortality: a reanalysis of the Steubenville data. Inhal. Toxicol. 7:35-44.

Morris, T. (1996) Regression analysis of daily mortality and air pollution using real and random data. Submitted as attachment to comments by American Petroleum Institute, Docket No. A-95-54, IV-D-2247. March 12, 1997.

NAPA (1991) National Acid Precipitation Assessment Program. Acidic Deposition: State of Science and Technology. Volume III. Report No. 24. 1991.

Neas LM, Dockery DW, Burge H, Koutrakis P, Speizer FE. 1996. Fungus spores, air pollutants and other determinants of peak expiratory flow rate in children. Am J Epidemiol 143:797-807

Neas L.M.; Dockery D.W.; Koutrakis P.; Tollerud, D.J.; Speizer F.E. (1995) The association of ambient air pollution with twice daily peak expiratory flow rate mesurements in children. Am. J. Epidemiol. 141:111-122.

Ostro, B.D.; Lipsett, M.J.; Wiener, M.B.; Selner, J.C. (1991) Asthmatic responses to airborne acid aerosols. Am J Public Health. 81:694-702.

Özkaynak, H. and Thurston, G.D. (1987). Associations between 1980 U.S. mortality rates and alternative measures of airborne particle concentration. Risk Analysis. 7(4):449-461.

Özkaynak, H.; Xue, J.; Severance, P.; Burnett, R.; Raizenne, M. (1994) Associations between daily mortality, ozone, and particulate air pollution in Toronto, Canada. Presented at: Colloquium on particulate air pollutant and human mortality and morbidity: program and abstracts; January, Irvine, CA: University of California Irvine, Air Pollution Health Effects Laboratory; p. P1-P13, report no. 94-02.

Özkaynak, H., Xue, J., Spengler, J., Wallace, L., Pellizzari, E, Jenkins, P. (1996a). Personal Exposure to Airborne Particles and Metals: Results from the Particulate Team Study in Riverside, California. Journal *of* Exposure Analysis and Environmental Epidemiology 6:(1) 57-78.

Özkaynak, H., Xue, J., Weker, R., Butler, D., Koutrakis, P. and Spengler, J. (1996b). The Particle TEAM (PTEAM) Study: Analysis of the Data. Final Report. Volume 3. Research Triangle Park, NC, EPA/600/R-95/098.

Özkaynak H, and Spengler, JD (1996) The Role of Outdoor Particulate Matter in Assessing Total Human Exposure. Chapter 4 in Particles in Our Air, Wilson R and Spengler, JD eds. Harvard University Press, Cambridge.

Pellizzari, E.D.; Thomas, K.W.; Clayton, C.A.; Whitmore, R.W.; Shores, R.C.; Zelon, H.S.; Peritt, R.L. (1993) Particle total exposure assessment methodology (PTEAM) Riverside, California Pilot Study - Volume I. U.S. Environmental Protection Agency, Atmospheric Research and Exposure Assessment Laboratory, Research Triangle Park, NC 27711, Report No. EPA/600/SR-93/050.

Perry, G.B.; Chai, H.; Dickey, D.W.; Jones, R.H.; Kinsman, R.A.; Morrill, C.G.; Spector, S.L.; Weiser, P.C. (1982) Effects of particulate air pollution on asthmatics. Am J Public Health. 73:50-56.

Peters A, Dockery DW, Heinrich J, Wichmann HE. 1997a. Short-term effects of particulate air pollution on respiratory morbidity in asthmatic children. Eur Respir J. 10:872-879

Peters A, Dockery DW, Heinrich J, Wichmann HE. 1997b. Medication use modifies the health effects of particulate sulfate air pollution in children with asthma. Environ Health Perspect 105: (in press)

Peters A, Wichmann HE, Tuch T, Heinrich J, Heyder J. 1997c. Respiratory effects are associated with the number of ultrafine particles. Am J Respir Crit Care Med 155: (in press)

Peters A, Doring A, Wichmann HE, Koenig W. 1997d. Increased plasma viscosity during the 1985 air pollution episode: a link to mortality? Lancet 349:1582-1587

Peters A, Goldstein IF, Beyer U, Franke K, Heinrich J, Dockery DW, Spengler JD, Wichmann HE. 1996. Acute health effects of exposure to high levels of air pollution in Eastern Europe. Am J Epidemiol 144:570-581

Ponce de Leon A, Anderson HR, Bland JM, Strachan DP, Bower J. 1996. Effects of air pollution on daily hospital admissions for respiratory disease in London between 1987-88 and 1991-92. J Epidemiol Comm Health 50(Suppl 1):S63-S70

Ponka A, Virtanen M. 1996. Asthma and ambient air pollution in Helsinki. J Epidemiol Comm Health 50(Suppl 1):S59-S62

Pope, C.A., III, Kalkstein, J.S. (1996) Synoptic weather modeling and estimates of the exposure-response relationship between daily mortality and particualte air pollution. Environ. Health. Perspect. 104:414-420.

Pope, C.A., III; Dockery, D.W.; Spengler, J.D.; Raizenne, M.E. (1991) Respiratory health and PM10 pollution: a daily time series analysis. Am Rev Resp Dis. 144:668-674.

Pope, C.A., III; Thun, M.J.; Namboori, M.M.; Dockery, D.W.; Evans, J.S.; Speizer, F.E.; Heath, D.W., Jr. (1995) Particulate air pollution as a predictor of mortality in a prospective study of U.S. adults. Am. J. Resp. Crit. Care Med. 151:669-674.

Pope, C.A., III; Dockery, D.W. (1992) Acute health effects of PM_{10} pollution on symptomatic and asymptomatic children. Am. Rev. Resp. Dis. 145:1123-1128.

Pope, C.A., III; Schwartz, J.; Ransom, M.R. (1992) Daily mortality and PM_{10} pollution in Utah Valley. Arch. Environ. Health. 47:211-217.

Price (1996) Letter from Dr. James H. Price, Chair Technical Subcommittee for Fine Particle Monitoring Clean Air Scientific Advisory Committee to Dr. George T. Wolff regarding Report of the CASAC Technical Subcommittee for Fine Particle Monitoring. August 1, 1996.

Pritchard, R.J., Ghio, A.J., Lehmann, J.R., Winsett, D.W., Tepper, J.S., Park, P., Gilmour, M.I., Dreher, K.L., Costa, D.L. (1996) Inhalation Toxicol., 8: 457-477.

Raizenne, M.; Neas, L.M.; Damokosh, A.I.; Dockery, D.W.; Spengler, J.D.; Koutrakis, P.; Ware, J.H.; Speizer, F.E. (1996) Health effects of acid aerosols on North American children: pulmonary function. Environ. Health. Perspect. 104:506-514.

Richmond, H. (1997) Memorandum to the File: Corrections to Risk Analysis Tables and Figures in PM Staff Paper. July 2, 1997.

Rodes, C.E., Holland, D.M., Purdue, L.J., and Rehme, K. (1985). A Field Comparison of PM_{10} Inlets at Four Locations, J. Air Waste Mgmt. Assoc., 35: 345-354.

Rojas-Bracho, L., Bahadori, T., Koutrakis, P., Suh, H. (1996) Personal, Indoor, and Outdoor Exposures to Particulate Matter of COPD Patients Living in Private Residences. Abstract K3.02, SRA/ISEA Annual Meeting, New Orleans, December 8 -12, 1996.

Romieu I, Meneses F, Ruiz S, Sienra JJ, Huerta J, White MC, Etzel RA. 1996. Effects of air pollution on the respiratory health of asthmatic children living in Mexico City. Am J Respir Crit Care Med 154:300-307

Roth, H.D. et al. (1997) Analysis of the association between total particulate matter and daily mortality in the Czech Republic: 1986-1994. Submitted as an attachment to comments of R.E. Wyzga, Docket No. A-95-54, IV-D-2329. March 11, 1997.

Roth, H.D., Li, Y. (1997) Analysis of the association between air pollutants with mortality and hospital admissions in Birmingham, Alabama: 1986-1990. Submitted as an attachment to comments by R.E. Wyzga, Docket No. A-95-54, IV-D-2672. March 11, 1997.

Saldiva, P.H.N.; Lichtenfels, A.J.F.C.; Paiva, P.S.O.; Barone, I.A.; Martins, M.A.; Massad, E.; Pereira, J.C.R.; Xavier, V.P.; Singer, J.M.; Bohm, G.M. (1994) Association between air pollution and mortality due to respiratory diseases in children in Sao Paolo, Brazil: a preliminary report. Environ Res. 65:218-225.

Saldiva, P.H.N.; Pope, C.A.,III.; Schwartz, J.; Dockery, D.W.; Lichtenfels, A.J.; Salge, J.M.; Barone, I.; Bohm, G.M. (1995) Air pollution and mortality in elderly people: a time-series study in Sao Paolo, Brazil. Arch Environ Health. 50:159-163.

Samet, J.M.; Zeger, S.L.; Kelsall, J.E.; Xu, J.; Kalkstein, L.S. (1996b) Weather, air pollution, and mortality in Philadelphia, 1973-1980: Report to the Health Effects Institute on Phase I.B of the Particle Epidemiology Evaluation Project. Health Effects Institute, Cambridge, MA.

Samet, J.M.; Zeger, S.L.; Berthane, K. (1995) The association of mortality and particulate air pollution. In: Particulate Air Pollution and Daily Mortality: Replication and Valication of Selected Studies (The Phase I.A Report of the Particle Epidemiology Evaluation Project) pp. 3-104. Health Effects Institute, Cambridge, MA.

Samet, J.M.; Zeger, S.L.; Kelsall, J.E.; Xu, J. (1996a) Air pollution and mortality in Philadelphia 1973-1988: Report to the Health Effects Institute on Phase I.B of the Particle Epidemiology Evaluation Project. Health Effects Institute, Cambridge, MA.

Samet, J.M., Stonehuerner, J., Reed, W., Devlin, R.B., Dailey, L.A., Kennedy, T.P., Bromberg, P.A., & Ghio, A.J. (1997) Disruption of protein tyrosine phosphate homeostasis in bronchial epithelial cells exposed to oil fly ash. Am. J. Physiol. 272 (Lung Cell. Mol. Physiol. 16): L426-L432.

Samet, J.M., Reed, W., Ghio, A.J., Devlin, R.B., Carter, J.D., Dailey, L.A., Bromberg, P.A., & Madden, M.C. (1996) Induction of prostaglandin H synthase 2 in human airway epithelial cells exposed to residual oil fly ash. Toxicol. Appl. Pharmacol., 141: 159-168.

Scarlett JF, Abbott KJ, Peacock JL, Strachan DP, Anderson HR. 1996. Acute effects of summer air pollution on respiratory function in primary school children in southern England. Thorax 51:1109-1114

Schouten JP, Vonk JM, de Graaf A. 1996. Short term effects of air pollution on emergency hospital admissions for respiratory disease: results of the APHEA project in two major cities in The Netherlands, 1977-89. J Epidemiol Comm Health 50(Suppl 1):S22-S29

Schlesinger, R.B., El-Fawal, H.A.N., Zelikoff, J.T., Gorczynski, J.E., McGovern, T., Nadziejko, C.E., Chen, L.C. (1994) Pulmonary effects of repeated episodic exposures to nitric acid vapor alone and in combination with ozone. Inhal Toxicol 6:21-41

Schwartz J, Spix C, Touloumi G, Bacharova L, Barumamdzadeh T, Le Tertre A, Piekarski T, Ponce de Leon A, Ponka A, Rossi G, Saez M, Schouten JP. 1996. Methodological issues in studies of air pollution and daily counts of deaths or hospital admissions. J Epidemiol Comm Health 50(Suppl 1):S3-S11

Schwartz, J,; Dockery, D.W.; and L.M. Neas. (1996). Is Daily Mortality Associated Specifically with Fine Particles? J Air & Waste Manag. Assoc. 46:927-939.

Schwartz, J. (1994a) Air pollution and hospital admissions for the elderly in Detroit, Michigan. Am J Resp Crit Care Med 159:648-655.

Schwartz, J. (1994b) Air pollution and hospital admissions for the elderly in Birmingham, Alabama. Am. J. Epidemiol. 139:589-598.

Schwartz, J.; Dockery, D.W.; Neas, L.M.; Wypij, D.; Ware, J.H.; Spengler, J.D.; Koutrakis, P.; Speizer, F.E.; Ferris, Jr., B.G. (1994) Acute effects of summer air pollution on respiratory symptom reporting in children. Am. J. Respir. Crit. Care. Med. 150:1234-1242.

Schwartz, J. (1993) Air pollution and daily mortality in Birmingham, Alabama. Am .J. Epidemiol. 137:1136-1147.

Schwartz, J; Dockery, D.W. (1992a) Increased mortality in Philadelphia associated with daily air pollution concentrations. Am. Rev. Resp. Dis. 145:600-604.

Schwartz, J.; Dockery, D.W. (1992b) Particulate air pollution and daily mortality in Steubenville, Ohio. Am. J. Epidemiol. 135:12-19

Silverman, F.; Hosein, H.R.; Corey, P,; Holton, S.; Tarlo, S.M. (1992) Effects of particulate matter exposure and medication use on asthmatics. Arch Environ Health 47:51-56.

Sisler, J.; Malm, W.; Molenar, J.; Gebhardt, K. (1996) Spatial and Seasonal Patterns and Long Term Variability of the Chemical Composition of the Haze in the U.S.: An Analysis of Data from the IMPROVE Network. Fort Collins, CO: Cooperative Institute for Research in the Atmosphere, Colorado State University. July 1996.

Spengler, J.D., Briggs, S.L.K., Ozkaynak, H. (1986) Relationships between TSP measurements and size-fractionated particle mass measurements in six cities participating in the Harvard air pollution health study. Report prepared for the U.S. EPA, Office of Air Quality Planning and Standards.

Spengler, J.D., Turner, W., Dockery, D.W. (1980) Comparison of hi-vol, dichotomous, and cyclone samplers in four U.S. cities. Presented at the 73rd Annual Meeting of the Air Pollution Control Association, Montreal, Canada. Paper 80-43.4.

Spix C, Wichmann H-E. 1996. Daily mortality and air pollutants: findings from Koln, Germany. J Epidemiol Comm Health 50(Suppl 1):S52-S58

Styer, P.; McMillan, N.; Gao, F.; Davis, J.; Sacks, J. (1995) Effect of outdoor airborne particulate matter on daily death counts. Environ Health Perspect. 103:490-497.

Su, W.-Y., Kodavanti, U.P., Jaskot, R.H., Costa, D.L., & Dreher, K.L. (1995) Temporal expression and cellular distribution of pulmonary fibronectin gene induction following exposure to an emission source particle. J. of Env. Pathol., Toxicol., and Oncol., 14(3&4): 215-225.

Suggs, J.C. and R.M. Burton. Spatial Characteristics of Inhalable Particles in the Philadelphia Metropolitan Area, J. APCA, 33, No. 7, pp. 688-691, July 198

Suh, H.H.; Allen G.A.; Koutrakis, P.; Burton, R.J. (1995). Spatial variation in acidic sulfate and ammonia concentrations within metropolitan Philadelphia, J Air Waste Manage Assoc, 45: 442-452.

Suh, H.H., Koutrakis, P, Spengler JD. (1993) Validation of personal exposure models for sulfate and aerosol strong acidity. J Air Waste Manage Assoc 43: 845-850.

Sunyer J, Castellsague J, Saez M, Tobias A, Anto JM. 1996. Air pollution and mortality in Barcelona. J Epidemiol Comm Health 50(Suppl 1):S76-S80

Tamura, K., Ando, M, Sagai, M., Matsumoto, Y. (1996) Estimation of levels of personal exposure to suspended particulate matter and nitrogen dioxide in Tokyo. Environ. Sci. (Tokyo) 6: in press.

Terashima, T., Wiggs, B., English, D., Hogg, J.C., van Eeden, S.F., (1997) Phagocytosis of small carbon particles (PM10) by alveolar macrophages stimulates the release of polymorphonuclear leukocytes from bone marrow. Am. J. Respir. Crit Care Med., 155: 1441-1447.

Thurston, G.D.; Ito, K.; Hayes, C.G.; Bates, D.V.; Lippmann, M. (1994) Respiratory hospital admissions and summertime haze air pollution in Toronto, Ontario: consideration of the role of acid aerosols. Environ. Res. 55:271-290.

Thurston, G.D.; Ito, K.; Kinney, P.L.; Lippmann, M. (1992) A multi-year study of air pollution and respiratory hospital admissions in three New York State Metropolitan Areas: results for 1988 and 1989 summers. J Expos Anal Environ Epidemiol. 2:429-450.

Touloumi G, Samoli E, Katsouyanni K. 1996. Daily mortality and "winter type" air pollution in Athens, Greece -- a time series analysis within the APHEA project. J Epidemiol Comm Health 50(Suppl 1):S47-S51

Touloumi, G.; Pocock, S.J.; Katsouyanni, K.; Trichopoulos, D. (1994) Short-term effects of air pollution on daily mortality in Athens: a time-series analysis. Int J Epidemiol. 23:957-967.

U.S. Environmental Protection Agency (1996a) Air Quality Criteria for Particulate Matter. Research Triangle Park, NC: National Center for Environmental Assessment. Office of Research and Development. April 12, 1996.

U.S. Environmental Protection Agency (1996b) Review of the National Ambient Air Quality Standards for Particulate Matter: Policy Assessment of Scientific and Technical Information -- OAQPS Staff Paper. Office of Air Quality Planning and Standards. Office of Air and Radiation. EPA report no. EPA-452\R-96-013. July 1996.

U.S. Environmental Protection Agency (1996c) Transcript of the Clean Air Scientific Advisory Committee's Review of the Particulate Matter Staff Paper Meetings held on May 6-7, 1996 in Chapel Hill, NC.

U.S. Environmental Protection Agency (1995) Office of Air and Radiation, Acid Rain Division. Acid Deposition Standard Feasibility Study Report to Congress. EPA report no. EPA 430-R-95-001a. October 1995.

U.S. Environmental Protection Agency (1993a) Office of Air Quality Planning and Standards Effects of the 1990 Clean Air Act Amendments on Visibility in Class I Areas: An EPA Report to Congress. Research Triangle Park, NC.

U.S. Environmental Protection Agency (1993b) Air Quality Criteria for Oxides of Nitrogen. Office of Research and Development. Research Triangle Park, NC: Office of Health and Environmental Assessment, Environmental Criteria and Assessment Office; EPA report no. EPA/600/8-91/049a-cF. August 1993.

U.S. Environmental Protection Agency (1989) An acid aerosols issue paper: health effects and aerometrics. Research Triangle Park, NC: Office of Health and Environmental Assessment, Environmental Criteria and Assessment Office; EPA report no. EPA-600/8-88-005F.

U.S. Environmental Protection Agency (1986) Second addendum to air quality criteria for particulate matter and sulfur oxides (1982): assessment of newly available health effects information. Research Triangle Park, NC: Office of Health and Environmental Assessment, Environmental Criteria and Assessment Office; EPA report no. EPA-600/8-86-020F.

U.S. Environmental Protection Agency (1982) Air quality criteria for particulate matter and sulfur oxides. Research Triangle Park, NC: Office of Health and Environmental Assessment, Environmental Criteria and Assessment Office; EPA report no. EPA-600/8-82-029a-cF.

Valberg, P.A. (1995) Comments before CASAC, December 14, 1995, Official Transcript, pages 152-157, 164 (ECAO-CD-92-0671, II.A.F.044).

Valdberg, P.A. (1997) Causality has not been demonstrated between outdoor levels of particulate matter (PM) and daily mortality and morbidity. In: Comments of the Engine Manufacturers Association on the proposed revisions to NAAQS for particulate matter and ozone. Docket No. A-95-54, IV-D-2328. March 11, 1997.

Vedal, S. (1997) Critical Review: Ambient Particles and Health: Lines that Divide. J Air Waste Manage Assoc. 47(5) 551-581.

Vigotti MA, Rossi G, Bisanti L, Zanobetti A, Schwartz J. 1996. Short term effects of urban air pollution on respiratory health in Milan, Italy, 1980-89. J Epidemiol Comm Health 50(Suppl 1):S71-S75

Ware, J.H.; Ferris, B.G., Jr.; Dockery, D.W.; Spengler, J.D.; Stram, D.O.; Speizer, F.E. (1986) Effects of ambient sulfur oxides and suspended particles on respiratory health of children. Am Rev. Resp. Dis. 133:834-842.

Watkinson, W. P.; Campen, M. J.; Costa, D. L. (1997) Cardiac arrhythmia induction after exposure to residual oil fly ash particles in a rodent model of pulmonary hypertension. Research Triangle Park, NC: U.S. Environmental Protection Agency, National Health and Environmental Effects Research Laboratory; submitted to J. Appl. Toxicol. and Fundam. Appl. Toxicol.

Wilson, WE and H.H. Suh. Fine Particles and Coarse Particles: Concentration Relationships Relevant to Epidemiologic Studies, accepted for publication in J Air Waste Manage Assoc, 1997.

Wojtyniak B, Piekarski T. 1996. Short term effect of air pollution on mortality in Polish urban populations - what is different? J Epidemiol Comm Health 50(Suppl 1):S36-S41

Wolfe, G.T. (1997) In Response to the PM Debate. Regulation, Winter 1997, p. 9

Wolff, G.T. (1996a) Letter from George T. Wolff, Chair, Clean Air Scientific Advisory Committee, to Administrator Carol M. Browner. Closure letter on draft Air Quality Criteria for Particulate Matter. March 15, 1996.

Wolff, G.T. (1996b) Letter from George T. Wolff, Chair, Clean Air Scientific Advisory Committee, to Administrator Carol M. Browner. Clsoure letter on draft OAQPS Staff Paper (Review of the National Ambient Air Quality Standards for Particulate Matter: policy Assessment of Scientific and Technical Information). June 13, 1996.

Woodruff, T.J.; Grillo, J.; Schoendorf, K.C.; 1997. The relationship between selected causes of postneonatal infant mortality and particulate air pollution in the United States. Environ. Health. Perspect. 105:(in press)

Wordley J, Walters S, Ayres JR. 1997. Short term variations in hospital admissions and mortality and particulate air pollution. Occup Environ Med 54:108-116

Wyzga, R.E.; Lipfert, F.W. (1995) Temperature-pollution interactions with daily mortality in Philadelphia. In: Particulate matter: health and regulatory issues: proceedings of an international specialty conference; April; Pittsburgh, PA. Air & Waste Management Association, Pittsburgh, PA (A&WMA Publication VIP-49)

Xu, X.; Gao, J.; Dockery, D.W.; Chen, Y. (1994) Air pollution and daily mortality in residential areas of Beijing, China. Arch Environ Health. 49:216-222.

Yoshimura, I. (1990) The effect of measurement error on the dose-response curve. Environ Health Perspect 87:173-178

Zmirou D, Barumandzadeh T, Balducci F, Ritter P, Laham G, Ghilardi J-P. 1006. Short term effects of air pollution on mortality in the city of Lyon, France, 1985-90. J Epidemiol Comm Health 50(Suppl 1):S30-S35