State of California

Air Resources Board

Mary D. Nichols, Chairman

Governor Arnold Schwarzenegger

Methodology for Estimating Premature Deaths Associated with Long-term Exposures to Fine Airborne Particulate Matter in California

DRAFT

Staff Report

December 7, 2009

California Environmental Protection Agency

Air Resources Board

California Environmental Protection Agency

Linda S. Adams, Secretary

Printed on Recycled Paper

This page is intentionally blank.

Methodology for Estimating Premature Deaths Associated with Long-term Exposures to Fine Airborne Particulate Matter in California

DRAFT
Staff Report

Project Coordinator and Lead Author

Hien T. Tran*

Contributing Authors

Álvaro Alvarado, Ph.D.
Cynthia Garcia
Nehzat Motallebi, Ph.D.
Lori Miyasato, Ph.D.
William Vance, Ph.D.

Scientific Advisors

Jonathan I. Levy, Sc.D., Harvard University
Bart Ostro, Ph.D., Office of Environmental Health Hazard Assessment, Cal EPA
Clive Arden Pope, III, Ph.D., Brigham Young University

Reviewers

Linda Tombras Smith, Ph.D., Chief, Health and Exposure Assessment Branch
Bart E.Croes, P.E., Chief, Research Division
Michael H. Scheible, Deputy Executive Officer

*NOTE: Report in identical form was published October 24, 2009. This author's Ph.D. title was removed after it was discovered it was not from an accredited institution. No other change was made to this report.

Peer Reviewers

Jeff Brook, Ph.D., Environment Canada
Mark D. Eisner, M.D., M.P.H., UC San Francisco
Richard C. Flagan, Ph.D., California Institute of Technology
Alan Hubbard, Ph.D., UC Berkeley
Joel Kaufman, M.D., M.P.H., University of Washington
Joel Schwartz, Ph.D., Harvard University

Acknowledgments

The staff thanks Donald McCubbin of Abt Associates, Inc. for assisting with BenMAP. Staff also appreciates the thorough review by Dr. Deborah Drechsler, the data analysis provided by Dr. Susan Gilbreath and Jeff Austin, and the technical assistance provided by Sarah Barnett, Vanessa Fontana, and Rebecca Boyer.

Disclaimer

This report has been reviewed by the staff of the Air Resources Board. Mention of trade names or commercial products does not constitute endorsement or recommendation for their use. To obtain this document in an alternative format, please contact the Air Resources Board ADA Coordinator at (916) 322-4505, TDD (916) 324-9531, or (800) 700-8326 for TDD calls from outside the Sacramento area.

Table of Contents

Executive Summary 1

I. Introduction and Background 2

II. Methodology 3

A. Summary of health studies on long-term PM exposures and premature death 3
B. U.S. EPA elicitation process 10
C. Applicability of U.S. EPA's expert elicitation results to California 20
D. Methodology for developing a concentration-response relationship 23
E. Methodology for estimating health impacts associated with PM exposures .. 26
F. Methodology for estimating ambient concentrations of PM from diesel-fueled engine emissions 29
G. Methodology for evaluating risk to small populations exposed to PM emissions from specific sources 31

III. Results

A. General relationship (relative risk) for use in California 32
B. Results on premature deaths associated with exposures to ambient PM 33
C. Results on premature deaths avoided by strategies designed to attain ambient air quality standards 35
D. Results on premature deaths associated with exposures to diesel PM exposures 36
E. Results on premature deaths associated with exposures to specific sources 37

IV. Discussion 38

V. Uncertainties and Limitations 39

VI. Conclusions 41

VII. References 43

APPENDICES

Appendix 1: (PM2.5 Exposures) A-3

Appendix 2: (Methodology for Estimating Health Impacts Avoided by Strategies Designed to Attain the Standards) A-9

Appendix 3: (Methodology for Estimating Ambient Concentrations of Particulate Matter from Diesel-Fueled Engine Emissions) A-13

Appendix 4: (Peer Review Process and Results) A-23

List of Figures

Figure 1: Expert uncertainty distributions for PM2.5-mortality coefficient, conditional on the existence of a causal relationship 15

Figure 2: Expert uncertainty distributions for PM2.5-mortality coefficient incorporating the experts' likelihood of a causal relationship 16

List of Tables

Table 1: Summary of key studies and relative risks on long-term exposures to PM and premature death 4

Table 2a: Key studies discussed by experts while answering conditioning questions.. 18

Table 2b: Key studies relied upon by experts in creating their C-R uncertainty distributions 19

Table 3: Percent change in mortality risk per 10 μg/m^3 increase in PM2.5 exposure.... 33

Table 4a: Annual premature deaths associated with ambient PM2.5 levels above 7 μg/m^3 34

Table 4b: Annual premature deaths associated with ambient PM2.5 levels above 2.5 μg/m^3 34

Table 4c: Annual premature deaths avoided by attainment of the national annual PM2.5 standard of 15 μg/m^3 35

Table 4d: Annual premature deaths avoided by attainment of the national annual PM2.5 standard of 12 μg/m^3 36

Table 5: Annual premature deaths associated with exposures to estimated primary diesel PM..... 37

Table 6: Annual premature deaths associated with PM2.5 from Goods Movement activities..... 38

Acronyms and Abbreviations

ACS	American Cancer Society
AHSMOG	The Adventist Health Study of Smog
ARB	California Air Resources Board
Board	California Air Resources Board
BS	Black Smoke or British Smoke
CARB	California Air Resources Board
CHS	Children's Health Study
C-R Function	Concentration-Response Function
CRPAQS	California Regional PM10/PM2.5 Air Quality Study
DPM	Diesel Particulate Matter
EI	Emission Inventory
EPA	United States Environmental Protection Agency
IMPROVE	Interagency Monitoring for Protected Visual Environments
NO	Nitric Oxide
NO_2	Nitrogen Dioxide
NO_x	Nitrogen Oxides
OEHHA	Office of Environmental Health Hazard Assessment
PM	Particulate Matter
PM2.5	Fine Particulate Matter; Particulate Matter 2.5 Micrometers in Diameter and Smaller
PM10	Particulate Matter 10 Micrometers in Diameter and Smaller
SA	Source Apportionment
SO_x	Sulfur Oxides
UFP	Ultrafine Particle
U.S. EPA	United States Environmental Protection Agency
VA	Veterans Administration

This page is intentionally blank.

Executive Summary

The California Air Resources Board (CARB) quantifies health impacts of exposure to particulate matter (PM) as part of the development of control measures for PM, including those for ports and goods movement. The methodology that CARB staff uses for quantifying premature death and other health impacts from PM exposure is based on a peer-reviewed methodology developed by the U.S. Environmental Protection Agency (EPA) for their risk assessments. This methodology is regularly updated by CARB staff as new epidemiological studies and other related studies are published that are relevant to California's health impacts analysis. This report discusses the results of staff's review of the recent scientific literature related to the mortality effects of exposure to fine PM (PM2.5) and presents recommendations for revisions to the current methodology.

In this report, the relative risk of premature death due to PM2.5 exposure was reevaluated based on all relevant scientific literature, and a new relative risk factor was developed. This new relative risk factor is a 10% increase in premature death per 10 $\mu g/m^3$ increase in PM2.5 exposures (uncertainty interval: 3% to 20%). Using this new factor, staff estimates diesel PM contributes to 3,900 (uncertainty interval 1,200 to 7,100) premature deaths, statewide on an annual basis. Staff also used a systematic approach for assessing the lowest level of PM2.5 that can be associated with premature death. Although the recent literature is consistent with a no-threshold model, no empirical evidence has been reported to date for an effect of exposure below 7 $\mu g/m^3$ in a general population. Staff therefore recommends that the cut-off be presented as a range of results. Using this approach, exposures to ambient PM2.5 can be associated with about 14,000 to 24,000 premature deaths statewide annually, with uncertainty ranging from 4,300 to 41,000 deaths.

The methodologies and results presented in this report have been endorsed by our scientific advisors, Dr. Jonathan Levy of Harvard University, Dr. Bart Ostro of the Office of Environmental Health Hazard Assessment, and Dr. Arden Pope of Brigham Young University. This report underwent an external peer review by experts selected through a process involving the University of California at Berkeley, Institute of the Environment. The results of the peer review process have been incorporated into this report.

I. Introduction and Background

In 2002, when CARB established a new ambient air quality standard for PM2.5 in collaboration with the Office of Environmental Health Hazard Assessment, we estimated the human health impacts of public exposures to PM levels above various levels, including the new standard (CARB 2002). The quantification of premature death from PM exposure used by CARB staff in previous analyses is based on a peer-reviewed methodology developed by the U.S. Environmental Protection Agency for their risk assessments (U.S. EPA 2004, 2005). The quantified death estimates play an important role in CARB's cost-benefits analysis of plans and regulations as they make up for the majority of the health valuation. For example, as part of the development of emission reduction plans and control measures for PM, CARB quantifies the health impacts of reducing population exposure to ambient PM that would result through the implementation of the proposed measures (CARB 2003a, 2003b, 2003c, 2004a, 2004b, 2004c).

In all of the recent analyses, including that for goods movement (CARB 2006), CARB has relied on the results from the American Cancer Society (ACS) study (Pope et al. 1995, 2002) to estimate the premature deaths. In U.S. EPA's regulatory impact analyses of recent years, including those on the Clean Air Interstate Rule in 2005, U.S. EPA continued to base the concentration-response function relating PM exposure to premature death on the published results of Pope et al. (2002). A concentration-response function relates changes in exposures to ambient concentrations of a pollutant to changes in an adverse health effect. However, several new epidemiological studies and other related studies have been published which may be relevant to California's health impacts analysis. These recent studies prompted CARB to consider updating the PM2.5 mortality relationship. For example, Jerrett et al. (2005) analyzed the data in the Los Angeles region, and Laden et al. (2006) performed an extended follow-up to the Harvard Six Cities study. In addition, intervention studies (Clancy et al. 2002) examining the effect of significant decreases in air pollution exposures show that the PM-mortality relationship can be larger than predicted by daily time-series studies (Samet et al. 2000). Also, clinical and toxicological studies (Chen et al. 2005) have emerged that suggest mechanisms by which PM exposure may contribute to the cardiovascular disease process, thus adding to the plausibility of the positive association between PM exposure and disease found in the long-term cohort studies.

Additional information comes from the U.S. EPA, which has elicited the opinions of twelve experts on the PM2.5-mortality relationship. Their opinions have been included in the latest regulatory impact analysis for the new national PM ambient air quality standard to characterize the uncertainty and range in the relationship[1], although Pope et al. (2002) results are still used in the primary analysis along with Laden et al. (2006).

At the April 20, 2006 Board meeting, staff presented the results of the goods movement

[1] http://www.epa.gov/ttn/ecas/regdata/RIAs/Chapter%205--Benefits.pdf

health impacts analysis (CARB 2006). Staff also informed the Board of plans to revise and improve the health impacts methodology by updating the health information that relates changes in PM2.5 exposures to premature death. This report is a product of this effort to update the methodology. In it, we summarize the health literature on the subject, interpret U.S. EPA's expert elicitation results, and explain how we apply these results to estimate the mortality impacts associated with Californians' exposures to ambient PM levels.

II. Methodology

The methodology presented in this report have been endorsed by our scientific advisors, Dr. Jonathan Levy of Harvard University, Dr. Bart Ostro of the Office of Environmental Health Hazard Assessment, and Dr. Arden Pope of Brigham Young University.

A. Summary of health studies on long-term PM exposures and premature death

The following is adapted with the authors' permission from the 2006 Critical Review in the Journal of Air and Waste Management Association by C. Arden Pope III and Douglas Dockery (Pope and Dockery 2006).

Daily time-series studies of acute exposures suggest short-term acute PM effects, but they provide little information about the degree of life shortening, pollution effects on longer-term premature death rates, or the role of pollution in inducing or accelerating the progress of chronic disease. As early as 1970, several analyses of pollution and premature death data reported that long-term average concentrations of PM2.5 or sulfate are associated with annual mortality rates across U.S. metropolitan areas. These population-based cross-sectional mortality rate studies were largely discounted by 1997 because of concern that they could not control for individual risk factors, such as cigarette smoking and body weight, which could potentially confound the air pollution effects. With regard to the premature death effects of long-term PM exposure, recent emphasis has been on prospective cohort studies that can control for individual differences in age, sex, smoking history, and other risk factors. However, since these studies require collecting information on large numbers of people and following them prospectively for long periods of time, conducting such studies can be costly, time consuming, and, therefore, much less common. A brief summary of results from these studies is presented in Table 1.

Below is a summary of the main long-term cohort studies published in the literature.

A.1 <u>Original Harvard Six Cities and ACS Studies</u>

In the mid-1990s, two cohort-based mortality studies had reported evidence of mortality effects of chronic exposure to fine particulate air pollution. The first study, often referred to as the Harvard Six Cities Study, reported on a 14- to 16-year prospective follow-up of

8,000 adults living in six U.S. cities (Dockery et al. 1993), representing a wide range of pollution exposure. The second study, referred to as the ACS (American Cancer Society) study, linked individual risk factor data from the ACS, Cancer Prevention Study II with national ambient air pollution data (Pope et al. 1995). The analysis included data from more than 500,000 adults who lived in 151 metropolitan areas and were followed prospectively from 1982 through 1989. About 50 metropolitan areas had PM and sulfate monitoring data. Both the Harvard Six Cities and the ACS cohort studies used Cox proportional hazard regression modeling to analyze survival times and to control for individual differences in age, sex, cigarette smoking, education levels, body mass index, and other individual risk factors. In both studies, cardiopulmonary mortality was significantly and most strongly associated with sulfate and PM2.5 concentrations.

Table 1: Summary of key studies and relative risks on long-term exposures to PM and premature death (Pope and Dockery 2006)

			Percent Increases in Relative Risk of Mortality (95% CI)		
Study	**Primary Source**	**Exposure Increment**	**All Cause**	**Cardiopulmonary**	**Lung Cancer**
Harvard Six Cities, original	Dockery et al. 1993	10 μg/m^3 PM2.5	13 (4.2, 23)	18 (6.0, 32)	18 (-11, 57)
Harvard Six Cities, HEI reanalysis	Krewski et al. 2000	10 μg/m^3 PM2.5	14 (5.4, 23)	19 (6.5, 33)	21 (-8.4, 60)
Harvard Six Cities, extended analysis	Laden et al. 2006	10 μg/m^3 PM2.5	16 (7, 26)	28 (13, 44)[a]	27 (-4, 69)
Harvard Six Cities, extended analysis between periods	Laden et al. 2006	10 μg/m^3 PM2.5	27 (5, 43)	31 (-1, 54)	6 (-57, 162)
ACS, original	Pope et al. 1995	10 μg/m^3 PM2.5	6.6 (3.5, 9.8)	12 (6.7,17)	1.2 (-8.7, 12)
ACS, HEI reanalysis	Krewski et al. 2000	10 μg/m^3 PM2.5	7.0 (3.9, 10)	12 (7.4, 17)	0.8 (-8.7, 11)
ACS, extended analysis	Pope et al. 2002 Pope et al. 2004	10 μg/m^3 PM2.5	6.2 (1.6, 11)	9.3 (3.3, 16) 12 (8, 15)[a]	13.5 (4.4, 23)
ACS adjusted using various education weighting schemes	Dockery et al. 1993 Pope et al. 2002 Krewski et al. 2000	10 μg/m^3 PM2.5	8–11	12–14	3–24
ACS intrametro Los Angeles	Jerrett et al. 2005	10 μg/m^3 PM2.5	17 (5, 30)	12 (_3, 30)	44 (-2, 211)
Postneonatal infant mortality, U.S	Woodruff et al. 1997	20 μg/m^3 PM10	8.0 (4, 14)	–	–
Postneonatal infant mortality, CA	Woodruff et al. 2006	10 μg/m^3 PM2.5	7.0 (-7, 24)	113 (12, 305)[c]	–
AHSMOG[b]	Abbey et al. 1999	20 μg/m^3 PM10	2.1 (-4.5, 9.2)	0.6 (-7.8, 10)	81 (14, 186)
AHSMOG, males only	McDonnell et al. 2000	10 μg/m^3 PM2.5	8.5 (-2.3, 21)	23 (-3, 55)	39 (-21, 150)
AHSMOG, females only	Chen et al. 2005	10 μg/m^3 PM2.5	–	42 (6, 90)[a]	–
Women's Health Initiative	Miller et al. 2004	10 μg/m^3 PM2.5	–	32 (1, 73)[a]	
Women's Health Initiative	Miller et al. 2007	10 μg/m^3 PM2.5	–	76 (25, 147)[a]	
VA, preliminary	Lipfert et al. 2000, 2003	10 μg/m^3 PM2.5	0.3 (NS)[d]	–	–
VA, extended	Lipfert et al. 2006	10 μg/m^3 PM2.5	15 (5, 26)[e]	–	–
11 CA counties, elderly	Enstrom 2005	10 μg/m^3 PM2.5	1 (-0.6, 2.6)	–	–
Netherlands	Hoek et al. 2002	10 μg/m^3 BS	17 (-24, 78)	34 (-32, 164)	–
Netherlands	Hoek et al. 2002	Near major road	41 (-6, 112)	95 (9, 251)	–
Netherlands	Beelen et al. 2008	10 μg/m^3 BS		22 (-1, 50)[c]	3 (-12, 20)
Netherlands	Beelen et al. 2008	10 μg/m^3 PM2.5		4 (-10, 21)[c]	6 (-18, 38)
Hamilton, Ontario, Canada	Finkelstein et al. 2004	Near major road	18 (2, 38)	–	–
French PAARC	Filleul et al. 2005	10 μg/m^3 BS	7 (3, 10)[f]	5 (-2,12)[f]	3 (-8,15)[f]
Cystic fibrosis	Goss et al. 2004	10 μg/m^3 PM2.5	32 (-9, 93)	–	–

[a]Cardiovascular only; [b]Pooled estimates for males and females; pollution associations were observed primarily in males and not females; [c]Respiratory only; [d]Reported to be nonsignificant by author; overall, effect estimates to various measure of particulate air pollution were highly unstable and not robust to selection of model and time windows; [e]Estimates from the single pollutant model and for 1989 –1996 follow-up; effect estimates are much smaller and statistically insignificant in an analysis restricted to counties with nitrogen dioxide data and for the 1997–2001 follow-up; furthermore, county-level traffic density is a strong predictor of survival and stronger than PM2.5 when included with PM2.5 in joint regressions; [f]Estimates when six monitors that were heavily influenced by local traffic sources were excluded; when data from all 24 monitors in all areas were used, no statistically significant associations between mortality and pollution were observed.

Although both the Harvard Six Cities and ACS studies used similar study designs and methods, these two studies had different strengths and limitations. The strengths of the Harvard Six Cities Study were its elegant and relatively balanced study design, the prospective collection of study-specific air pollution data, and the ability to present the core results in a straightforward graphical format. On the other hand, the primary limitations of the Harvard Six Cities Study were the small number of subjects from a small number of study areas (that is, exposures) in the Eastern United States. In contrast, the major strength of the ACS study was the large number of participants and cities distributed across the entire United States. The primary limitation of the ACS was the lack of planned, prospective collection of study-specific air pollution and health data, and the reliance on limited, separately collected subject and pollution data. Nonetheless, the ACS study provided a test of the hypotheses generated from the Harvard Six Cities Study in an independently collected dataset. Therefore, these two studies were considered complementary.

A.2 Reanalyses and Extended Analyses of Harvard Six Cities and ACS Studies

In the mid-1990s, the Harvard Six Cities and the ACS prospective cohort studies provided compelling evidence of mortality effects from long-term fine particulate air pollution (Dockery et al. 1993, Pope et al. 1995). Nevertheless, these two studies were controversial. Subsequently, the data quality, accessibility, analytic methods, and validity of these studies came under intense scrutiny when the U.S. EPA considered them in the effort to revise the PM ambient air quality standards. There were serious constraints and concerns regarding the dissemination of confidential information and the intellectual property rights of the original investigators and their supporting institutions. In 1997, the investigators of the two studies agreed to provide the data for an intensive reanalysis by an independent research team under Health Effects Institute (HEI) oversight, management, sponsorship, and under conditions that assured the confidentiality of the information on individual study participants. The reanalysis included: (1) a quality assurance audit of the data, (2) a replication and validation of the originally reported results, and (3) sensitivity analyses to evaluate the robustness of the original findings. The reanalysis (Krewski et al. 2000, 2004) reported that the data were "generally of high quality" and that the results originally reported could be reproduced and validated. The data audit and validation efforts revealed some data and analytic issues that required some tuning. However, the adjusted results did not differ substantively from the original findings. The reanalysis demonstrated the robustness of the PM-mortality risk estimates to many alternative model specifications. Further, the reanalysis team also made a number of innovative methodological contributions that not only demonstrated the robustness of the PM-mortality results but substantially contributed to subsequent analyses. In the reanalysis, persons with higher educational attainment were found to have lower relative risks of premature death associated with PM2.5 in both studies.

Further extended analyses of the ACS cohort (Pope et al. 2002, 2004) included more than twice the follow-up time (more than 16 years) and approximately triple the number of deaths. The mortality associations with fine particulate and sulfur oxide pollution

persisted and were robust to control for individual risk factors including age, sex, race, smoking, education, marital status, body mass index, alcohol use, occupational exposures, and diet and the incorporation of both random effects and nonparametric spatial smoothing components. There was no evidence that the PM-mortality associations were due to regional or other spatial differences that were not controlled in the analysis. These analyses also evaluated associations with expanded pollution data, including gaseous co-pollutant data and new PM2.5 data. Elevated premature death risks were most strongly associated with measures of PM2.5 and sulfur oxide pollution. Coarse particles and gaseous pollutants, except for sulfur dioxide (SO_2), were generally not significantly associated with elevated premature death risk.

Jerrett et al. (2005a) assessed air pollution associations of the 23,000 subjects in the ACS cohort who lived in the metropolitan Los Angeles area. PM-mortality associations were estimated based on PM2.5 measures from 23 monitoring sites interpolated to 267 residential zip code centroids for 2000, and health data analyzed for the period between 1982 and 2000. Cox proportional hazards regression models controlled for age, sex, race, smoking, education, marital status, diet, alcohol use, occupational exposures, and body mass. In addition, because variations in exposure to air pollution within a city may correlate with socioeconomic gradients that influence health and susceptibility to environmental exposures, zip code-level ecological variables were used to control for potential "contextual neighborhood confounding" (Jerrett et al. 2003, 2005b). The premature death associations with the intra-metropolitan PM2.5 concentrations were generally larger than those observed previously in the ACS cohort across metropolitan areas. However, the associated confidence intervals were also wider than those previously reported in the ACS national cohort studies. Nonetheless, such results corroborate the Harvard Six Cities results (Dockery et al. 1993), making the possibility of a greater effect than observed in the full ACS cohort more plausible.

A recent analysis of the Harvard Six Cities cohort by Laden et al. (2006) extended the mortality follow-up for eight more years with approximately twice the number of deaths. PM2.5 concentrations for the extended follow-up years were estimated from PM10 and visibility measures. PM2.5-mortality associations, similar to those found in the original analysis, were observed for all-cause, cardiovascular, and lung cancer mortality. However, PM2.5 concentrations were substantially lower for the extended follow-up period than they were for the original analysis, especially for two of the most polluted cities. Reductions in PM2.5 concentrations were associated with reduced premature death risk and were largest in the cities with the largest declines in PM2.5 concentrations. The authors note that, "these findings suggest that mortality effects of long-term air pollution may be at least partially reversible over periods of a decade." Further, it is noteworthy that the authors observed a substantial decrease in premature death risk corresponding to the decrease in PM2.5 concentrations between the two periods.

A.3 Other Independent Studies

The Adventist Health Study of Smog (AHSMOG)

The Adventist Health Study of Smog (AHSMOG) cohort study related air pollution to 1977–1992 mortality in more than 6000 non-smoking adults living in California, predominantly from San Diego, Los Angeles, and San Francisco (Abbey et al. 1999). All-cause mortality, nonmalignant respiratory mortality, and lung cancer mortality were significantly associated with ambient PM10 concentrations in males but not in females. Cardiopulmonary disease mortality was not significantly associated with PM10 in either males or females. This study did not have direct measures of PM2.5 but relied on TSP and PM10 data. In a follow-up analysis (McDonnell et al. 2000), visibility data were used to estimate PM2.5 exposures of a subset of males who lived near an airport. All-cause, lung cancer, and nonmalignant respiratory disease (either as the underlying or a contributing cause) were more strongly associated with PM2.5 than with PM10. In a recent analysis of the AHSMOG cohort, fatal coronary heart disease was significantly associated with PM among females but not among males (Chen et al. 2005).

Women's Health Initiative

The association between long-term PM2.5 exposure and first cardiovascular events (fatal and nonfatal) were explored in the Women's Health Initiative Observational Study (Miller et al. 2004, Miller et al. 2007). Based on measurements from the nearest monitor, air pollution exposures were estimated for about 66,000 post-menopausal women without prior cardiovascular disease in 36 metropolitan areas from 1994 to 1998. After adjusting for age, smoking, and various other risk factors, PM2.5 exposures were found to be significantly associated with increases in nonfatal cardiovascular and fatal cardiovascular events, including premature death from cardiovascular disease. The risk of death from exposure to PM2.5 was greater than nonfatal cardiovascular events. The hazard ratio estimated from this study was also larger than mortality estimates from other studies. The authors suggest that the larger hazard ratio may be due to efforts to reduce misclassification of outcomes and exposures. It may also be possible that the effects of PM2.5 may be greater in women than men. Because this study investigated the association between long-term PM2.5 exposure and first cardiovascular events, it is unlikely that the effects are limited only to women who are already ill.

Veterans Administration (VA)

Lipfert et al. (2000, 2003) assessed the association of total mortality and air pollution in a prospective cohort of about 50,000 middle-aged, hypertensive, male patients from 32 Veterans Administration (VA) clinics followed for about 21 years. The cohort had a disproportionately large number of current or former smokers (81%) and African-Americans (35%) relative to the U.S. population or to other cohorts that have been used to study air pollution. Air pollution exposures were estimated by averaging air pollution data for participants' county of residence at the time of entrance into the cohort. Only analyses of total mortality were reported. In addition to considering mortality and average exposures over the entire follow-up period, three sequential mortality periods and four exposure periods were defined and included in various analyses. Lipfert et al. (2006a) extended the follow-up of the VA cohort and focused on traffic density as the measure of environmental exposure. It was suggested that traffic density was a more "significant and robust predictor of survival in this cohort" than PM2.5. However, of the

various measures of ambient air pollution, PM2.5 was most strongly correlated with traffic density (r = 0.50). In single pollutant models, PM2.5 was associated with mortality risk resulting in risk estimates comparable to other cohorts. These results were also confirmed in another analysis by Lipfert et al. (2006b) examining PM2.5 constituents and related air quality variables as predictors of survival. Overall, in the VA analyses, effect estimates to various measures of PM were unstable and not robust to model selection, time windows used, or various other analytic decisions. It was difficult, based on the preliminary results presented, to make conclusive statistical inferences regarding PM-mortality associations.

Eleven California Counties

Enstrom (2005) reported an analysis of about 36,000 elderly males and females in 11 California counties followed between 1973 and 2002. Countywide PM2.5 concentrations were estimated from outdoor ambient monitoring for the time period 1979–1983. For approximately the first half of the follow-up period (1973–1983) and for the time period approximately concurrent with PM2.5 monitoring, a small PM2.5-mortality association was observed. No PM2.5-mortality risk associations were observed for the later follow-up (1983–2002). For the entire follow-up period, only a small statistically insignificant association was observed. When 1979-93 pollution and mortality date were examined, a statistically significant association was observed.

Netherlands Pilot Study

In a pilot study, Hoek et al. (2002) evaluated the associations between premature death and PM based on a random sample of 5000 participants in the Netherlands Cohort Study on Diet and Cancer, originally 55 to 69 years of age and followed for more than 8 years. Although the effect estimates were not very precise, the adjusted risk of cardiopulmonary mortality was nearly double for individuals who lived within 100 meters of a freeway or within 50 meters of a major urban road. Based on residential location of participants and interpolation of pollution data from the Netherlands' national air pollution monitoring network, average background concentrations of black smoke ([BS] or British smoke measured by optical densities or light absorbance of filters used to gather PM from the air) for the first 4 years of follow-up were estimated. Background plus local traffic-related black smoke exposures were estimated by adding to the background concentration a quantitative estimate of living near a major road. Cardiopulmonary mortality was associated with estimates of exposure to black smoke, and the association was nearly doubled when local traffic-related sources of black smoke in addition to background concentrations were modeled.

A more recent study on the same Cohort, Beelen et al. (2008), reinforces the conclusions of the pilot study. The authors found a positive association between traffic intensity on the nearest roadway to the subject's residence and death rate. They also confirmed the link between interpolated BS concentrations and cardiopulmonary mortality. While the associations between pollutants and mortality in this study were not statistically significant, the authors' methodology was very careful, and their results lend convincing support to the link between premature death and PM.

Hamilton, Ontario, Canada

In an exploration of the relationship between proximity to traffic air pollution and premature death observed in the Netherlands study, an analysis using a cohort of 5,228 persons greater than 40 years of age living in Hamilton, Ontario, Canada, was conducted (Finkelstein et al. 2004). Somewhat higher mortality risks were observed for individuals who lived within 100 meters of a highway or within 50 meters of a major road.

Air Pollution and Chronic Respiratory Diseases (PAARC) Survey in France

Filleul et al. (2005) reported an analysis of about 14,000 adults who resided in 24 areas from seven French cities as part of the Air Pollution and Chronic Respiratory Diseases (PAARC) survey. Participants were enrolled in 1974, and a 25-year mortality follow-up was conducted. Ambient air pollution monitoring for total suspended particulates, black smoke, nitrogen dioxide, and nitric oxide was conducted for three years in each of the 24 study areas. When survival analysis was conducted using data from all 24 monitors in all of the areas, no statistically significant associations between mortality and pollution were observed. However, when the six monitors that were heavily influenced by local traffic sources were excluded, non-accidental mortality was significantly associated with all four measures of pollution, including black smoke. In addition to PM, mortality was associated with nitric oxide. Nitric oxide concentrations were also significantly associated with mortality risk in a cohort of Norwegian men (Nafstad et al. 2004), but no measure of PM was available.

Cystic Fibrosis Foundation

A unique study of the effects of ambient air pollution was conducted utilizing a cohort of 20,000 patients more than 6 years old who were enrolled in the U.S.-based Cystic Fibrosis Foundation National Patient Registry in 1999 and 2000 (Goss et al. 2004). Annual average air pollution exposures were estimated by linking fixed-site ambient monitoring data with resident zip code. A positive, but not statistically significant, association between PM2.5 and premature death was observed. PM2.5 was associated with statistically significant declines in lung function (FEV1) and an increase in the odds of two or more pulmonary exacerbations.

Postneonatal Infants

Woodruff et al. (1997) reported the results of an analysis of postneonatal infant mortality (deaths after one month of age and before one year of age determined from the U.S. National Center for Health Statistics birth and death records) for about 4 million infants in 86 U.S. metropolitan areas between 1989 and 1991 linked with U.S. EPA-collected PM10 data. Postneonatal infant mortality was compared with levels of PM10 concentrations during the 2 months after birth, controlling for maternal race, maternal education, marital status, month of birth, maternal smoking during pregnancy, and ambient temperatures. Postneonatal infant mortality for all causes, respiratory causes

and sudden infant death syndrome (SIDS) were associated with particulate air pollution. Woodruff et al. (2006) also linked monitored PM2.5 to infants who were born in California in 1999 and 2000 and who lived within 5 miles of a monitor, matching 788 postneonatal deaths to 3,089 survivors. Each 10 $\mu g/m^3$ increase in PM2.5 was associated with a near doubling of the risk of postneonatal death because of respiratory causes and a statistically insignificant increase for death from all causes.

A.4 Summary

Cohort studies generally apply proportional hazards models controlling for many individual-level risk factors (such as body mass index, smoking, alcohol use, occupational exposures, age/race, etc. and ecologic factors) before air pollution is considered. Many of the above studies also correct for spatial autocorrelation to avoid misinterpretation of results.

Nonetheless, evaluating which studies to consider in assessing the public health impacts of air pollution is a difficult task. As recommended by both the National Research Council (2002) and the Science Advisory Board (U.S EPA 2004), the U.S. EPA elicited experts for their assessment of the literature and opinion on the most appropriate concentration-response function relating premature death to long-term exposures to PM2.5. This process asked experts to review all available studies to derive the plausible range of values that describe the PM2.5-mortality relationship. These studies included not only the cohort studies described above but also intervention studies which show stronger effects compared to time-series or cohort studies. Also included were toxicological and clinical studies which suggest the mechanisms by which PM exposures can contribute to the cardiovascular disease process, thus adding to the plausibility of the positive association between exposures and disease found in the long-term cohort studies.

B. U.S. EPA elicitation process

In this section, we adapt a report by U.S. EPA's contractor, Industrial Economics (2006) to describe the U.S. EPA's expert elicitation. Similar information has been published in Environmental Science and Technology (Roman et al. 2008).

In its 2002 report to Congress titled *Estimating the Public Health Benefits of Proposed Air Pollution Regulations*, the National Research Council (2002) recommended that a better characterization of the uncertainty be performed for regulatory impact analyses, including estimating premature death associated with exposures to PM2.5 levels.[2] As a result, U.S. EPA convened a panel of twelve experts to assess the reduction in premature death in the adult U.S. population resulting from a long-term reduction in annual average PM2.5. Our proposed methodology makes use of results from the panel's report. In their assessment, the experts considered all published literature on

[2] http://www.nap.edu/catalog/10511.html

the subject.[3]

B.1 Selection of experts

The twelve experts participating in the study were selected through a two-part peer nomination process and included experts in epidemiology, toxicology, and medicine. The peer nomination process was designed to obtain a balanced set of views and serves to minimize the influence of Industrial Economics and U.S. EPA on expert selection.

The first phase of the expert selection process was designed to select nine experts. The initial decision to include nine experts was based on several factors, including: 1) a literature search that found most of the elicitation studies conducted to date (60 percent) use panels of six to eight experts, and 90 percent use panels of 11 or fewer experts (Walker, 2004); 2) it was deemed that nine experts would provide a balanced set of views on this topic; 3) the pilot study conducted in 2004 was criticized for the small panel size of five experts (IE 2004); 4) government agencies are required to undergo an Information Collection Request process for the Paperwork Reduction Act if information is collected from more than nine individuals; and 5) resource and time requirements increase with each additional expert.

While this process featured a good acceptance rate and yielded nine experts, the panel exhibited less diversity in expertise than originally anticipated in design, with most experts being epidemiologists. In an effort to increase representation of the biological, medical, and toxicological disciplines, a second phase of selections was conducted. U.S. EPA sought additional nominations of experts in these fields based on nominations provided by the Health Effects Institute (HEI). The general criteria for nominations were the same as for the first part of the selection process (Holmstead 2005).

The following twelve individuals made up the panel of experts:

- Doug Dockery, Ph.D., Professor of Environmental Epidemiology
  Department of Environmental Health, Harvard School of Public Health
- Kaz Ito, Ph.D., Assistant Professor of Environmental Medicine
  New York University of Medicine
- Daniel Krewski, Ph.D., Director
  R. Samuel McLaughlin Centre for Population Health Risk Assessment
  University of Ottawa
- Nino Kuenzli, M.D., Ph.D., Associate Professor
  Department of Preventive Medicine
  University of Southern California Keck School of Medicine
- Morton Lippmann, Ph.D., Professor and Director of Aerosol Research
  Laboratory, New York University School of Medicine
- Joe Mauderly, DVM, Vice President and Senior Scientist
  Lovelace Respiratory Research Institute

[3] http://www.epa.gov/ttn/ecas/benefits.html

- Bart Ostro, Ph.D., Chief
  Air Pollution Epidemiology Unit,
  California Environmental Protection Agency Office of Environmental Health Hazard Assessment
- C. Arden Pope, III, Ph.D., Professor of Economics
  Brigham Young University
- Richard Schlesinger, Ph.D., Biology and Health Sciences
  Pace University
- Joel Schwartz, Ph.D., Professor of Environmental Health
  Department of Environmental Health, Harvard School of Public Health
- George Thurston, Ph.D., New York University of Medicine,
- Mark Utell, M.D., Professor of Medicine and Environmental Medicine
  University of Rochester School of Medicine and Dentistry

B.2 Elicitation process

A "briefing book" binder was sent to all experts at least two weeks in advance of their interview (IE 2006). The purpose of the briefing book was to provide experts with a baseline set of materials to assist them in preparing for their elicitation interview; however, experts were free to consider other materials not included in the briefing book. The briefing book contained the following materials: the elicitation interview protocol; a CD containing over 150 relevant papers and compendia, searchable both alphabetically and by topic area; a set of background information pages with recent U.S. data on air quality, health status, population demographics, and other topics that may factor into the experts' probabilistic judgments; and background materials, including a document describing factors to consider when providing probability judgments in order to avoid potential sources of bias, and an excerpt from the National Research Council (2002) report on estimating public health benefits of proposed air rules.

The pre-elicitation workshop was designed to introduce the project, provide background information to the panel on expert judgment and the elicitation process, and to foster discussion about the key evidence available to answer the questions posed by the study. The key evidence includes not only the main studies on long-term exposures to PM and mortality but also short-term time-series studies, toxicological studies, intervention studies, and other studies.

Each elicitation interview lasted approximately eight hours and covered both qualitative and quantitative questions. The qualitative questions probed experts' beliefs concerning key evidence and critical sources of uncertainty and were intended to make the conceptual basis for their quantitative judgments explicit. These questions covered topics such as potential biological mechanisms linking PM2.5 exposures with premature death; key scientific evidence on the magnitude of the PM-mortality relationship; sources of potential error or bias in epidemiological results; the likelihood of a causal relationship between PM2.5 and premature death; and the shape of the concentration-response (C-R) function. The main quantitative question asked each expert to provide a probabilistic distribution for the average expected decrease in U.S. annual, adult, all-cause mortality associated with a 1 $\mu g/m^3$ decrease in annual average PM2.5 levels.

In addressing this question, the experts first specified a functional form for the PM2.5 mortality C-R function and then developed an uncertainty distribution for the slope of that function (the mortality impact per unit change in annual average PM2.5), taking into account the evidence and judgments discussed during the qualitative questions.

When answering the main quantitative question, each expert was instructed to consider that the total mortality effect of a 1 μg/m^3 decrease in ambient annual average PM2.5 may reflect reductions in both short-term peak and long-term average exposures to PM2.5. Each expert was asked to aggregate the effects of both types of changes in his answers. Each expert was given the option to integrate their judgments about the likelihood of a causal relationship and/or threshold in the C-R function into his distribution or to provide a distribution "conditional on" one or both of these factors. The interviewers asked each expert to characterize his distribution by assigning values to fixed percentiles (5th, 25th, 50th, 75th, 95th). To assist experts in the elicitation process, the interviewers provided real-time feedback during the interviews in the form of graphs and example calculations, using spreadsheet tools and Internet teleconferencing. During the interviews, experts were able to view their responses plotted onto a distribution using a software interface. They then adjusted their estimates until the distribution represented the views they expressed during the day-long interview.

B.3 Results of U.S. EPA's elicitation

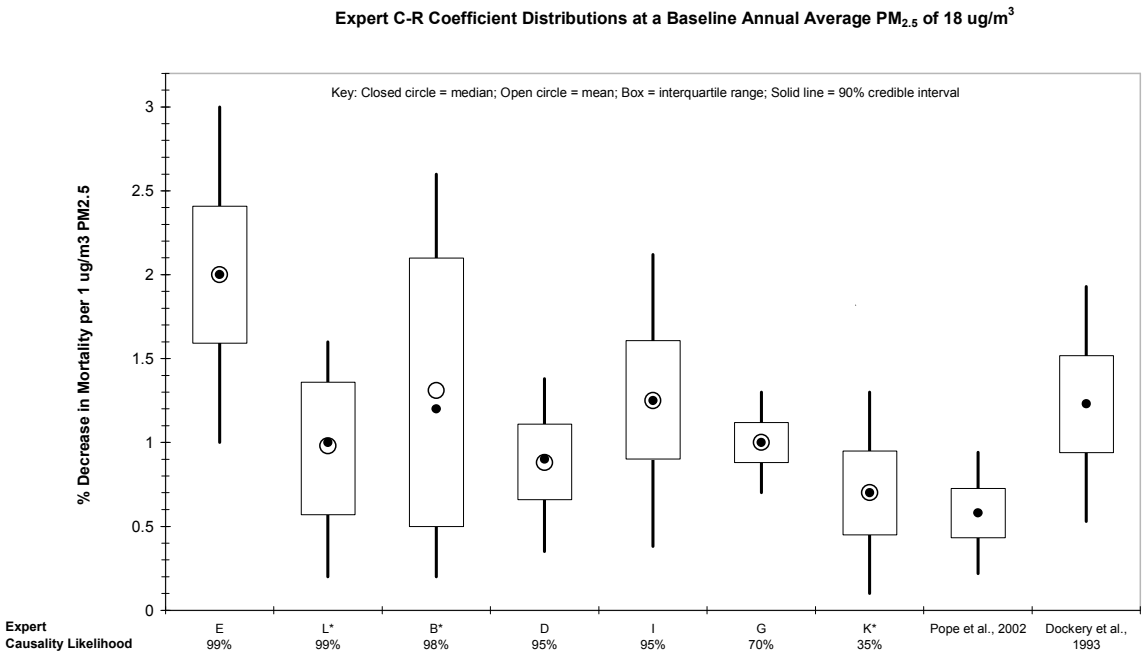
Figures 1 and 2 display the responses of the experts to the main quantitative elicitation question. The distributions provided by each expert, identified by the letters A through L, are depicted as box and whisker plots with the solid circle symbol showing the median (50th percentile); the open circle showing the mean; the box defining the interquartile range (bounded by the 25th and 75th percentiles); and the ends of the "whiskers" defining each expert's 5th and 95th percentiles.

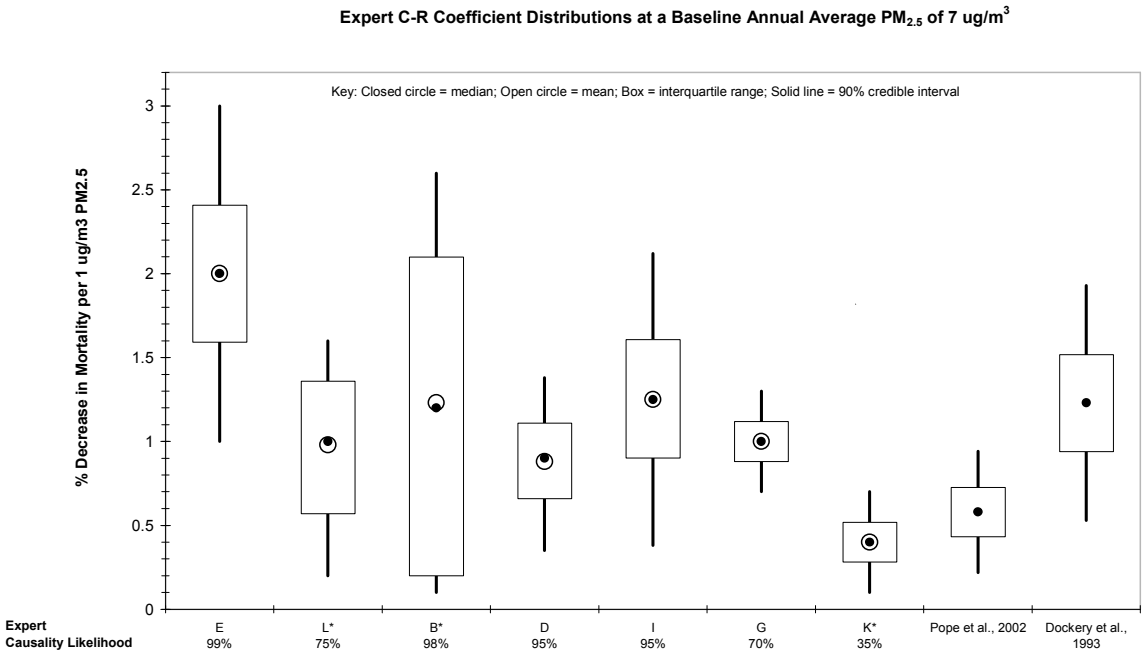
Each expert's stated best estimate of the likelihood of a causal relationship between PM2.5 and premature death is shown on the x-axis and the experts are arrayed in order of decreasing certainty of causality. Figure 1 displays the distributions for the experts who chose to provide a distribution conditional on the existence of a causal relationship between PM2.5 and premature death. Figure 2 shows the distributions for the group who chose to integrate their judgments about the likelihood of causality directly into their distribution. Each figure displays the expert distributions for two different PM2.5 levels, 18 μg/m^3 and 7 μg/m^3, to observe the implications of four experts' (B, F, K, and L) assumptions about nonlinearities in the C-R function and about differing degrees of uncertainty in the slope of the function across specific ranges of PM. Also, as a point of reference for the results, we include box plots of two epidemiologic studies often used in U.S. EPA benefit analyses (Pope et al. 2002, Dockery et al. 1993).

Among the experts who provided distributions that were conditional on the existence of a causal relationship (Figure 1), median estimates ranged from a 0.4 to 2.0 percent decrease in annual, adult, all-cause mortality risk per 1 μg/m^3 decrease in annual average PM2.5 exposures. Similarly, among the experts who directly incorporated their

views on the likelihood of a causal relationship into their distributions (Figure 2), the median estimates also ranged from a 0.7 to 1.6 percent decrease in annual, adult, all-cause mortality risk per 1 $\mu g/m^3$ decrease in annual-average PM2.5 exposures.

Figure 1: Expert uncertainty distributions for PM2.5-mortality coefficient, conditional on the existence of a causal relationship (IE 2006)





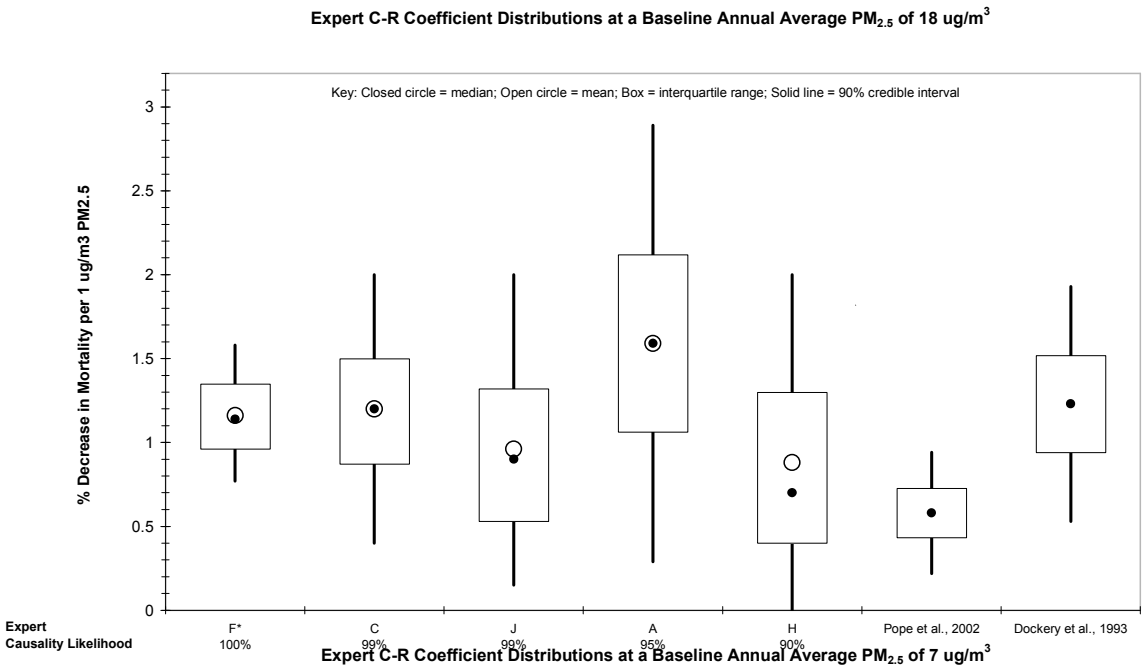
Note: Box plots represent distributions as provided by the experts to the elicitation team. When asked, experts in this group preferred to give conditional distributions and keep their probabilistic judgment about the likelihood of a causal or non-causal relationship separate.

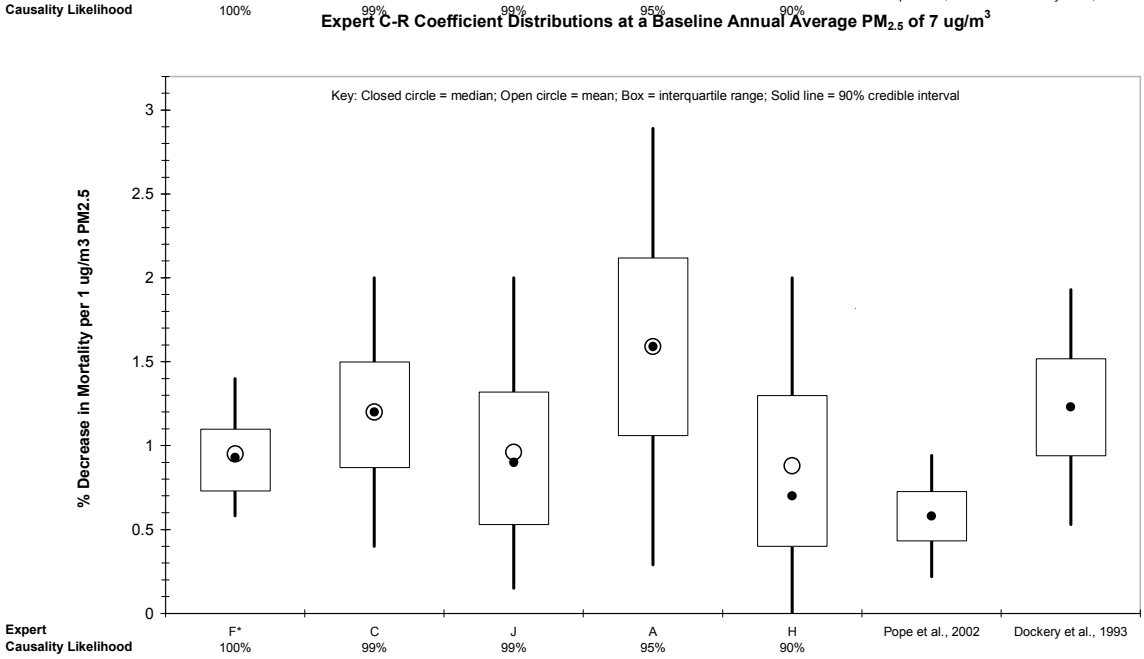
*Experts' C-R coefficient distribution changes between 7 and 18 $\mu g/m^3$.

Expert K specified a threshold (not shown).

Expert L provided two different likelihoods of causality for his C-R coefficient distributions at 7 and 18 $\mu g/m^3$, although his distribution appears in the same location in both graphs.

Figure 2: Expert uncertainty distributions for PM2.5-mortality coefficient incorporating the experts' likelihood of a causal relationship (IE 2006)





Note: Box plots represent distributions as provided by the experts to the elicitation team. When asked, experts in this group preferred to give distributions that incorporate their likelihood that the $PM_{2.5}$ mortality association may be non-causal

*Experts' C-R coefficient distribution changes between 7 and 18 $\mu g/m^3$.

Certain observations and conclusions can be drawn from these plots and from the experts' responses to the qualitative questions:

- Experts in this study tended to be confident that PM2.5 exposure can cause premature death. Ten of twelve experts believed that the likelihood of a causal relationship was 90 percent or higher. The remaining two experts gave causal probabilities of 35 and 70 percent. Recent research in both epidemiology (e.g., Jerrett et al. 2005, Laden et al. 2006) and toxicology (e.g., Sun et al. 2005) significantly contributed to experts' confidence.
- Only one of twelve experts explicitly incorporated a threshold into his C-R function. The rest believed there was a lack of empirical and/or theoretical support for a population threshold. However, three other experts gave differing effect estimate distributions above and below some cut-off concentration. The adjustments these experts made to median estimates and/or uncertainty at lower PM2.5 concentrations were modest.
- Experts relied upon a core set of cohort epidemiology studies to derive their quantitative estimates, mainly those associated with the ACS and Six Cities cohorts. The Six Cities results tended to be weighted more highly by experts in this study than in the pilot study. The greater emphasis on Six Cities appeared to result from corroborating evidence in the recent Six Cities follow-up (Laden et al., 2006) and from concerns about potential exposure misclassification issues and/or effect modification in the ACS cohort (see below). Expert K indicated that he was 50 percent sure that a threshold existed. If there were a threshold, he thought that there was an 80 percent chance that it would be less than or equal to 5 $\mu g/m^3$, and a 20 percent chance that it would fall between 5 and 10 $\mu g/m^3$. See Table 2a and 2b for a listing of core studies used by the experts.
- Although the quantitative question asked experts to consider mortality changes due both to short-term and long-term PM2.5 exposures, all experts based their median effect estimates on effects due to long-term exposures. Short-term exposure effects were sometimes used to derive lower-bound effect estimates.
- Confounding of epidemiological results tended to be a minor concern for most experts. Only one of twelve experts expressed substantial concern about confounding as a source of error in the key literature on PM2.5 and premature death.
- Experts' concerns regarding potential negative bias in the ACS main study results due to effect modification (see Pope and Dockery 2006) and/or exposure misclassification (Jerrett et al. 2005; Willis et al. 2003; and Mallick et al. 2002) led many experts to adjust the published results upwards when considering the percentiles of their distribution.
- A sensitivity analysis conducted using a simplified benefits analysis (IE 2006) demonstrated that no individual expert's distribution of effect estimates had more than a plus or minus 8 percent impact on an overall, pooled distribution of effects. The influence of individual experts appeared symmetrically distributed.

Table 2a: Key studies discussed by experts while answering conditioning questions (IE 2006)

	Six Cities (Dockery et al., 1993; Krewski et al., 2000; Laden et al., 2006)	ACS (Pope et al., 1995, 2002 & 2004; Krewski et al., 2000)	ACS LA Reanalysis (Jerrett et al., 2003 & 2005	Utah Valley (Pope et al., 1989, 1991, 1996; Ghio et al., 2004)	Netherlands Cohort Study (Hoek et al., 2002)	AHSMOG (Abbey et al., 1991 & 1999; McDonnell et al., 2000; Chen et al., 2005)	Veteran's (Lipfert et al., 2000, 2003 & 2006)	Dublin Study (Clancy et al., 2002)	Elderly Californians Study (Enstrom et al., 2005)	Hong Kong Study (Hedley et al., 2002)	APHEA[1]	NMMAPS (Samet et. al., 2000)	Canadian Time-Series Studies (Burnett et al., 2000 & 2003)	Finkelstein et al., 2004	MESA Cohort[2]	Mallick et al., 2002	Willis et al., 2002	Filluel et al., 2005	Woodruff et al., 1997	Women's Health Initiative[2]
Expert A	✓	✓	✓	✓	✓			✓			✓	✓			✓					
Expert B	✓	✓	✓			✓	✓				✓	✓	✓			✓				
Expert C	✓	✓	✓	✓				✓		✓										
Expert D	✓	✓	✓	✓			✓		✓											
Expert E	✓	✓	✓		✓												✓			
Expert F	✓	✓	✓																	
Expert G	✓	✓		✓						✓										
Expert H	✓	✓				✓	✓													
Expert I	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓							✓		
Expert J	✓	✓	✓		✓	✓	✓		✓					✓					✓	✓
Expert K	✓	✓	✓					✓												
Expert L	✓	✓	✓		✓	✓			✓											
Total:	12	12	10	5	5	5	5	4	4	3	3	2	1	1	1	1	1	1	1	1

[1] The Air Pollution and Health - A European Approach (APHEA) includes a large group of studies. For full list of papers, please consult http://airnet.iras.uu.nl/products/reports_and_annexes/APHEA/APHEA_publications.pdf.

[2] Study not yet published at the time of the interview.

Table 2b: Key studies relied upon by experts in creating their C-R uncertainty distributions (IE 2006).

	ACS (Pope et al., 2002)	ACS LA Reanalysis (Jerrett et al., 2005)	Six Cities (Dockery et al., 1993)	Six Cities (Laden et al., 2006 (Cross-Sectional))	ACS (Pope et al., 1995)	Netherlands Cohort Study (Hoek et al., 2002)	Six Cities (Laden et al., 2006 (Change estimate))	Mallick et al., 2002	Willis et al., 2002	NMMAPS (Samet et al., 2000)	Women's Health Initiative[2]	AHSMOG (Abbey et al., 1991 & 1999; McDonnell et al., 2000; Chen et al., 2005)
Expert A	○	⊙	○	⊙			○			○		
Expert B	⊙	⊙			⊙ ○			⊙				
Expert C	⊙	⊙	⊙									
Expert D	⊙		⊙									
Expert E	⊙	⊙		⊙		⊙	⊙		⊙			
Expert F	○	○	⊙	⊙								
Expert G	⊙ ○		⊙ ○									
Expert H	⊙	⊙ ○	○									
Expert I	⊙	⊙		⊙								
Expert J[1]	⊙	⊙ ○	⊙ ○			O				○	○	○
Expert K	⊙		⊙ ○									
Expert L	○	⊙ ○		○						○		
Total ⊙:	9	8	6	4	1	1	1	1	1	0	0	0
Total ○:	4	4	5	1	1	1	1	0	0	3	1	1

⊙ = Expert used the study to inform the median of his C-R coefficient distribution(s).

○ = Expert used the study to inform the uncertainty of his C-R coefficient distribution(s).

[1] Expert J also cited the following short-term studies as support for his uncertainty: Levy et al., 2000; Steib et al., 2002; Anderson et al., 2005; Ostro et al., 2005; Schwartz et al., 1996, Klemm et al., 2000; Burnett et al., 2003).

[2] Study not yet published at the time of the interview.

B.4 U.S. EPA's peer review process

Six reviewers were asked to participate in the peer review of U.S. EPA's elicitation. They include:

- Douglas Crawford-Brown, Ph.D.
  Director, Institute for the Enviroment
  Professor, Environmental Sciences and Engineering
  University of North Carolina at Chapel Hill
- John S. Evans, Sc.D.
  Senior Lecturer on Environmental Science
  Harvard School of Public Health
- Granger Morgan, Ph.D.
  Lord Chair Professor in Engineering
  Carnegie Mellon University
- D. Warner North, Ph.D.
  Department of Management Science and Engineering
  Stanford University
- David Stieb, Ph.D.
  Air Health Effects Division,
  Health Canada
- Thomas S. Wallsten, Ph.D.
  Department of Psychology
  University of Maryland at College Park

Overall, the reviewers unanimously agreed that U.S. EPA conducted a high quality expert elicitation. The elicitation follows best practices and can serve as a model of good practice for expert elicitations in a variety of agency-wide settings. The reviewers agree that the elicitation protocol provides a reliable basis for eliciting the probabilistic distributions of uncertainty in the PM2.5 C-R relationship[4].

C. Applicability of U.S. EPA's expert elicitation results to California

The experts' judgments on the PM2.5-mortality relationship apply to regulatory impact analyses at the national scale. To fully understand how such results are applicable to California, it is helpful to discuss the strengths and weaknesses of the studies cited by the experts and evaluate how applicable they are in California.

The studies described in Section II.A provide significant evidence regarding the influence of PM2.5 on premature death. However, only a subset of these studies may be suitable for developing a relative risk applicable to general populations for use in regulatory impact analyses. While the relative risk in premature death per unit change in PM2.5 long-term exposures is derived from a formal expert elicitation protocol, as described in Section II.B, by highlighting the strengths and weaknesses of the various studies from the perspective of relative risk derivation, we can better interpret the expert

[4] http://www.epa.gov/ttn/ecas/regdata/Uncertainty/pm_ee_peer_review_summary.pdf

elicitation output.

One key factor in choosing an appropriate study is the generalizability of the study population. As our objective is to derive a relative risk applicable to the general population of California, it is important to use studies that have a similar at-risk population. This criterion would eliminate direct application of studies like the Washington University-EPRI Veterans Cohort (Lipfert et al. 2000, 2003, 2006), which focused on male military veterans under treatment for hypertension, with 81 percent current or former smokers. Similarly, the Adventist Health Study on Smog (AHSMOG) (Abbey et al. 1999, McDonnell et al. 2000, Chen et al. 2005) focused on non-Hispanic white Seventh-Day Adventists who were nonsmokers. In addition, studies on infant mortality (Woodruff 1997, 2006) do not directly address long-term exposures to PM2.5; hence, they do not apply to our assessment. It is important to recognize that the inability to utilize these studies directly to develop general population relative risks does not mean that they are invalid, nor does it mean that these studies did not influence the judgments of the experts within the expert elicitation. Findings regarding the effect of PM2.5 on populations either with a greater or lesser collection of risk factors than the general population are informative, but cannot directly provide a relative risk applicable to the general population of California.

Other criteria that can be applied involve utilizing only studies that measured the pollutant of interest (PM2.5) and the health outcome of interest (all-cause mortality). Thus, while studies like Miller et al. (2004) and Chen et al. (2005) may be more interpretable by focusing on cardiovascular risk (an outcome for which there is extensive evidence supporting biological plausibility), if the aim is to develop a relative risk factor for all-cause mortality, these studies cannot be used directly. Similarly, studies that use an alternative measure of particulate matter like black smoke (Filleul et al. 2005) or proximity to a major road (Beelen et al. 2008, Hoek et al. 2003, Finkelstein et al. 2004) provide insight about the effects of motor vehicle-related particulate matter on premature death but cannot directly inform PM2.5 relative risk. In addition, the AHSMOG study also cannot be used directly, for it did not have direct measurement of PM2.5 but relied on TSP and PM10 data.

Other important screening criteria include a desire for geographic appropriateness. This does not necessarily mean that only studies in California can be used for risk evaluations in California, but it means that significant factors that vary geographically should be addressed. This can occur at multiple levels. For example, a study in a developing country may not be directly applicable to the U.S., due to differences in age distributions, underlying disease patterns, pollutant composition, standard of health care, and many other factors. Within the U.S., regional differences could occur if the composition of PM2.5 differed significantly and more/less toxic agents could be identified, or if concentration-exposure relationships differed significantly (i.e., due to differences in air conditioning prevalence). While there are some noticeable differences between California and other states in terms of climate and concentrations of PM constituents, there is little evidence for California's relative risk to be differentiated from the U.S. average. More explicitly, there is not adequate evidence at present regarding the quantitative differential toxicity of different particle constituents, and national and

regional information about exposure-concentration differentials, to make any formal adjustments.

National-scale epidemiological studies addressing short-term effects of PM exposures using time-series analyses do not demonstrate an appreciable difference between California and other states or regions in relative risks. For example, in a publication on 91 U.S. cities addressed by the National Mortality Morbidity Air Pollution Study, Dominici et al. (2005) showed that the southern California relative risk was slightly higher than the national average, while that of the Northwest (which included northern California as well as Oregon, Washington) was slightly lower than the national average. A simple average of the southern California and Northwest relative risks gives a value almost identical to the national average. A recent publication investigating PM2.5 mortality in 27 large communities around the U.S. (Franklin et al. 2007) found that the C-R function was above the national average for San Diego and Sacramento but below the national average and insignificant for Riverside and Los Angeles. It should be noted that the cohort study by Jerrett et al. (2005) did find a statistically significant effect for the Los Angeles metropolitan area, once exposure was estimated with more geographic precision. Thus, the available evidence does not provide any rationale for excluding relative risks derived from studies across the U.S. to California.

In addition, studies used in developing a relative risk for use in quantifying public health impacts should ideally have controlled for co-pollutants and other potentially significant confounders, should have undergone extensive sensitivity analyses, and been validated through multiple measures (i.e., detailed quality assurance/quality control, re-analyses by multiple investigators). These represent standard quality criteria for including studies in any meta-analyses; they also serve to guide us in choosing studies for California's risk assessments as well. In this regard, the Enstrom (2005) study of elderly Californians neither adequately controlled for smoking nor adjusted for exposure to environmental tobacco smoke, two factors that could significantly alter the effect of PM exposures on premature death. Further, exposure misclassification is another issue of concern. In Entrom's study, PM2.5 was assigned on the basis of data from just a few monitoring sites and at times on very few measurements (Brunekreef 2006). No discussion was provided as to the representativeness of sites; it is surprising, for instance, that Kern County ranked higher than Los Angeles in terms of PM exposures. Yet another issue is the long time passed since enrollment (1959) and follow-up (1973-2002), which must have been associated with many changes in diet, smoking, occupation, etc., factors for which the authors could not adequately control.

Based on the above criteria, the primary evidence for PM2.5 mortality C-R functions comes from multiple analyses from the Harvard Six Cities study (Dockery et al. 1993, Krewski et al. 2000, Laden et al. 2006) and the ACS cohort study (Pope et al. 1995, Krewski et al. 2000, Pope et al. 2002, Pope et al. 2004, Jerrett et al. 2005). Each of these studies addresses all-cause mortality associated with PM2.5 from a general population cohort, and each has undergone extensive peer review and re-analysis. In spite of the strengths, there are some limitations of each study. Namely, the Six Cities study focused on only white adults in six cities in the eastern half of the U.S., with resulting concerns for generalizability and for statistical power. The ACS study

addressed these concerns by considering a larger number of subjects and a more expansive geographic coverage, although some population representativeness issues remained due to the recruitment approach for the ACS Cancer Prevention Study II. There are also concerns that the retrospective exposure assessment (using existing monitors) may have contributed exposure misclassification, a point potentially supported by the greater C-R function in Jerrett et al. (2005) relative to earlier publications. Regardless, these studies fulfill all other criteria and can be used as a basis to develop a new relative risk for regulatory impact analyses in California. As can be seen in the discussion in Section II.B, the experts recruited by U.S. EPA relied heavily on these studies to develop their probability distributions of the PM2.5-mortality relationship.

In summary, it is appropriate to rely on the U.S. EPA's experts' judgments for California's specific risk assessments. Both the ACS national study by Pope et al. (1995, 2002), which includes California counties, and the ACS sub-cohort study in Los Angeles (Jerrett et al. 2005) heavily influenced the experts' evaluations. Although the Harvard Six Cities studies do not include California, the range in PM levels observed in the six cities reflect those measured in California, and the analysis by Jerrett et al. (2005) produced results similar to those found in the Harvard Six Cities studies. Thus, it is justifiable to use Harvard Six Cities studies for California. Furthermore, time-series studies like NMMAPS show the PM-mortality relationship holds for broad geographic regions, including California (Dominici et al. 2005). Hence, it is appropriate to rely on U.S. EPA's expert elicitation results in developing a new relationship between premature death and long-term PM exposures for use in California.

D. Methodology for developing a concentration-response relationship

While the expert elicitation protocol yields significant insight regarding the strength of current scientific evidence and the range of C-R functions supported by experts in the field, some caution is necessary in interpreting a pooled estimate or the collective opinion of the panel. Some researchers (Morgan and Henrion 1990) assert that, if the range of expert opinions is significant enough to have major consequences for the outcome of the analysis, the opinions should generally not be combined to produce an "average" result. The empirical evidence seems to indicate good agreement among most experts regarding the appropriate C-R function, in which case any pooling approach would yield similar estimates, but there are some important differences that may be masked or exaggerated by a combined estimate.

If a pooled estimate is needed for a given policy application, as is the case here, there are a few basic approaches that could be used. The simplest approach is to average the individual assessments, or similarly, to use inverse-variance weighted averages. While this has the benefit of simplicity, this approach presumes that all experts are equally well-calibrated in their abilities to construct confidence intervals, which is not likely the case. Many expert elicitation applications use a series of calibration exercises, utilizing questions for which the answer is known or knowable, to determine the ability of experts to characterize uncertainty. This ability is characterized by calibration (i.e., 5 percent of estimates are outside of a 95 percent confidence interval) and

informativeness (confidence intervals are not excessively large).

Within U.S. EPA's expert elicitation, no calibration exercise was done, so we do not have the ability to construct individual weights beyond the reported confidence intervals. Thus, it is potentially most interpretable to examine the range of estimates provided and determine a central estimate and low/high estimate, without conducting a formal statistical pooling of estimates. Among measures of central tendency, the median is the statistic least influenced by outlying observations. With that in mind, staff chose the median to represent the point of central tendency among each expert's distribution of point estimates. The median of the experts' medians is then considered to be the overall estimate of central tendency for the PM-mortality relationship. We also used the medians of the experts' 5th and 95th percentiles as the lower and the upper bound of the credible range, respectively. Consequently, the credible range can be treated as a 90% uncertainty interval around the estimate of the PM-mortality relationship.

D.1 Sensitivity Analysis

Simple averaging of experts' distributions can be used to corroborate the above assessments, with sensitivity analyses on the relative weights used to determine the robustness of the pooled estimate. In addition, results will be compared against pooling empirical study results. Later, we demonstrate that alternative approaches for deriving the central, low, and high estimates yield similar results to the approach CARB staff has chosen. Below is a detailed discussion of these alternative approaches.

Developing a credible range of the PM-mortality relationship based on a wide range of evidence on the subject is without doubt challenging. We demonstrate the robustness of our chosen range by considering several alternative ways to interpret the data and arrive at other plausible C-R functions. These include:

1. Pooling three studies, Pope et al. (2002), Laden et al. (2006), and Jerrett et al. (2005) using equal weight — to treat the results from three studies equally. Note that since Jerrett's analysis uses a subset of the ACS cohort analyzed by Pope et al., it is technically incorrect to pool the non-independent results. However, for the purpose of demonstrating the robustness of the approach CARB staff has chosen, results are presented in this report.
2. Pooling Pope et al. (2002), Laden et al. (2006) and Jerrett et al. (2005) using inverse-variance weighting — to give more weight to studies with tighter confidence bounds than those with wider confidence bounds.

The remaining four alternative analyses rely on random effects pooling, of which a detailed discussion follows.

3. Pooling Pope et al. (2002), Laden et al. (2006), and Jerrett et al. (2005) using random effects in BenMAP[5].
4. Pooling Pope et al. (2002) and Laden et al. (2006) using random effects in BenMAP.
5. Pooling all 12 expert distributions using random effects in BenMAP

[5] http://www.epa.gov/air/benmap/download.html

6. Pooling 10 expert distributions (without experts E & H, who provided the highest and lowest estimates among the twelve experts). This analysis will assess the impact of outlying opinions using random effects in BenMAP.

A common method for weighting estimates involves using their variances. The variance takes into account both the consistency of data and the sample size used to obtain the estimate, two key factors that influence the reliability of results. The exact way in which variances are used to weight the estimates from different studies in a pooled estimate depends on the underlying model.

The fixed effects model assumes that there is a single true concentration-response relationship and therefore a single true value for the parameter in question. For example, in our discussion, the parameter would be the relative risk. Differences among parameters reported by different studies are therefore simply the result of sampling error. That is, each reported relative risk is an estimate of the *same underlying parameter.* The certainty of an estimate is reflected in its variance (the larger the variance, the less certain the estimate). Pooling that assumes a fixed effects model therefore weights each estimate under consideration in proportion to the *inverse* of the variance. This means that estimates with small variances (i.e., estimates with relatively little uncertainty surrounding them) receive large weights, and those with large variances receive small weights.

The estimate produced by pooling based on a fixed effects model, then, is just a weighted average of the estimates from the studies being considered, with the weights as defined to be equal, as in scenario (1) above, or inverse-variance, as in scenario (2). An alternative to the fixed effects model is the random effects model, which allows the possibility that the estimated relative risks from the different studies may in fact be estimates of *different* parameters, rather than just different estimates of a single underlying parameter. In studies of the effects of PM on premature death, for example, if the level of wood burning varies among study locations the underlying relationship between mortality and PM may be different from one study location to another. If wood burning associated with cold weather causes individuals to stay inside more on days with high PM (likely to occur during the winter in California), then the mortality risk may be lower in areas with high prevalence of wood burning. As such, one would expect the true value of the relative risk in cities with low wood burning prevalence to be greater than the true value of the relative risk in cities with high wood burning prevalence. This would violate the assumption of the fixed effects model.

Embedded in BenMAP is a procedure for testing whether it is appropriate to base the pooling on the random effects model (vs. the fixed effects model). If the evidence does not support the fixed effects model, then the random effects model is assumed, allowing the possibility that each study is estimating a different relative risk. The weights used in a pooling based on the random effects model must take into account not only the within-study variances (used in a meta-analysis based on the fixed effects model) but the between-study variance as well. The weighting scheme used in a pooling based on the random effects model is basically the same as that used if a fixed effects model is

assumed, but the variances used in the calculations are different. This is because a fixed effects model assumes that the variability among the estimates from different studies is due only to sampling error (i.e., each study is thought of as representing just another sample from the same underlying population), while the random effects model assumes that there is not only sampling error associated with each study, but that there is also *between-study* variability — each study is estimating a different underlying beta coefficient. Therefore, the sum of the within-study variance and the between-study variance yields an overall variance estimate. U.S. EPA's report[6] provides a more detailed discussion of this weighting scheme.

Once a concentration-response function relating changes in PM exposures to premature death is derived, one can estimate the health impacts.

E. Methodology for estimating health impacts associated with PM exposures

In this section, we discuss the methodology developed to estimate the health impacts associated with PM exposures above a predetermined level. This methodology is consistent with that used in CARB's staff report on the PM ambient air quality standard (CARB, 2002). The major modification to that methodology is calculating impacts at an annual level for three years, then averaging the results, rather than averaging exposure estimates over three years and then calculating health impacts. This modification is an improvement over the previous methodology since the annual concentrations (not three-year average concentrations) are used to address the average-annual PM impact, and averaging over three years would yield results that are more representative of the current situation than just using one year of data. Detailed discussions of each step follow.

STEP 1: Obtain PM concentrations for all sites in California

The observed PM2.5 concentrations are obtained for years 2004, 2005, and 2006. In addition to the routine monitoring network, data from the IMPROVE (Interagency Monitoring for Protected Visual Environments) are included in the analysis. See Appendix 1 for a description of these special monitoring data. Annual averages of quarterly means are calculated for each site for consistency with the national and state definition of the PM standard attainment designations.

STEP 2: Estimate PM concentration per census tract

The concentration per census tract is estimated using the ambient annual average PM2.5 concentrations measured at monitoring sites. This step is done with BenMAP[7], a software program developed by the U.S. EPA for estimating and mapping health impacts associated with air pollution. BenMAP interpolates PM concentrations using nearby monitored values with the inverse distance weighted squared method.

[6] http://www.epa.gov/ttn/ecas/regdata/RIAs/Appendix%20H--Additional%20Details%20on%20Benefits%20Methodologies.pdf#search='epa%20benmap%20random%20effects%20pooling'

[7] http://www.epa.gov/air/benmap/download.html

The interpolation is confined to a 50-kilometer radius, with the weight assigned to each nearby monitored PM value as the inverse square of the distance from the monitor to the location of interpolation. In some areas of California, there may be no monitoring information within 50 kilometers. In these cases, the concentration that will be assigned will be from the closest monitor, regardless of the distance. The end result is a smooth contour surface of PM values throughout the entire state. The interpolated value is then assigned to each census tract center. This step is performed for each of the three years.

STEP 3: Estimate mortality impact

The concentration-response functions are applied to calculate mortality impacts due to long-term changes in PM exposure, using county-specific baseline incidence rates from the Center for Disease Control[8].

For log-linear functions, the health impact is

$\Delta Y = -Y_O [\exp(-\beta * \Delta PM - 1)] * pop$, where

Y_O = baseline mortality rates, which include all-case deaths for the population over age 30. We used the morality rate for the year 2005 to calculate health impacts for years 2004, 2005 and 2006.

β = beta coefficient derived from the relative risk of epidemiologic study results.

ΔPM = the difference between the estimated ambient PM concentration and a level below which we expect no PM-related mortality or cut-off level.

pop = population age 30 or above in each census block, from US Census for each year (2004-2006).

Note that the baseline mortality rate and population are available for various subgroups (age 30-34, 35-44, 45-54, 55-64, 65-74, 75-84, 85+). The health impact is actually calculated for each subgroup at the census track level. After each change in health impacts is calculated for each census track, we sum across the results for an air basin or for the entire state. Heath impacts are calculated for each year; they are then averaged over three years to reduce the influence of any year with unusual meteorology on the overall results.

E.1 Cut-off Level

This section describes ARB's consideration of a cut-off level or level below which we expect no PM-related mortality. Recent evidence suggests that exposures to low PM2.5 levels may lead to adverse health impacts (Schwartz et al. 2002, Kappos et al. 2004, de Kok et al. 2006, Miller et al. 2007). In addition, most of the long-term exposure studies

[8] http://wonder.cdc.gov/mortSQL.html

that examined the shape of the C-R function failed to demonstrate a flattening of the function at lower levels; linearity could not be rejected based on statistical tests (Krewski et al., 2000, Pope et al. 2002, Schwartz et al., 2008). Finally, many daily time-series mortality studies include concentrations very close to background levels (Ostro et al. 2006, Schwartz et al. 2002, Schwartz et al. 1996). For these reasons, we assessed the likelihood of a threshold by reviewing the scientific literature on this issue and by inferring from the conclusions of the U.S. EPA's expert elicitation.

As part of the protocol in the U.S. EPA's expert elicitation, the experts were asked for their individual judgment regarding whether a threshold exists in the PM2.5-mortality function. The purpose was to assess expert judgments regarding theory and evidential support for a population threshold (i.e., the concentration below which no member of the study population would experience an increased risk of death). From a theoretical and conceptual standpoint, all experts generally believed that while a threshold may exist at the individual level, there was no evidence of a population-based threshold. Specifically, eleven of the twelve experts discounted the idea of a population threshold in the C-R function on a theoretical and/or empirical basis. Seven of the experts favored epidemiological studies as ideally the best means of addressing the population threshold issue, suggesting this approach is best to evaluate the full range of susceptible individuals at environmentally relevant exposure levels. However, those who favored epidemiologic studies generally acknowledged that definitive studies addressing thresholds would be difficult or impossible to conduct since they would need to include a very large and diverse population with wide variation in exposure and a long follow-up period. The following is a discussion of three alternatives for a threshold level: 7, 2.5, and 5 $\mu g/m^3$.

<u>Cut-Off Level of 7 $\mu g/m^3$</u>. As discussed above, in the U.S. EPA's expert elicitation, only one of twelve experts thought the shape of concentration-response function may change at a level at or below 7 $\mu g/m^3$, suggesting that this level may serve as a possible threshold. The level of 7 $\mu g/m^3$ also happens to be the lowest concentration observed in the American Cancer Society study (Pope 2002). In this large cohort study, Pope et al. (2002) provided empirical evidence that exposures to PM2.5 levels as low as 7 $\mu g/m^3$ can be associated with premature death. Also, since the ACS study is the largest cohort study conducted to date, it would be reasonable to use 7 $\mu g/m^3$ as a presumed cut-off level for calculating PM2.5-related mortality. Thus, there is direct empirical evidence that some effects are likely to occur down to this level, although based on the limited data at these low concentrations, the uncertainty is greater than the uncertainty for PM levels in the middle range of the distribution (between 10 and 18 $\mu g/m^3$).

<u>Cut-Off Level of 2.5 $\mu g/m^3$</u>. A second alternative is to select the background level for PM2.5 as the cut-off level, which addresses all impacts due to anthropogenic PM exposures. In California, the background PM2.5 level is 2.5 $\mu g/m^3$ (Motallebi et al. 2003). As discussed above, there is no empirical evidence for long-term mortality effects at levels below 7 $\mu g/m^3$. Thus, quantifying human health impacts due to exposures at levels below 7 $\mu g/m^3$ would be the result of personal judgment and

inference from the available data on long-term studies. In the Women's Health Initiative Study (Miller et al. 2007), the investigators found significant relationships between long-term exposure to PM2.5 and the incidence of cardiovascular events at levels lower than 7 $\mu g/m^3$. However, due to the subpopulation of older women being addressed in this study, we could not justify using the results for a general population.

Cut-Off Level of 5 $\mu g/m^3$. During the review of the document, the peer reviewers were asked to consider the cut-off level in addressing premature death associated with PM2.5 exposures. The reviewers recognized that selecting a cut-off level involves professional judgment due to limited empirical evidence in the low PM2.5 range. The consensus of the peer review panel was that a cut-off level of 4 to 5 $\mu g/m^3$ was reasonable based on the lowest observed short-term levels associated with mortality (Ostro et al. 2006, Schwartz et al. 2002, Schwartz et al. 1996).

Staff Recommendation. While empirical evidence indicates that mortality can be associated with long-term exposure to PM2.5 levels as low as 7 $\mu g/m^3$, the consensus of the peer reviewers is that effects likely occur below this level. However, choosing a specific value for a threshold of effect is necessarily a matter of individual judgment, due to the lack of long-term data at low ambient concentrations of PM2.5. Assuming that the probability of effects between 7 $\mu g/m^3$ and 2.5 $\mu g/m^3$ (background) is uniform, staff recommends that the cut-off level be expressed as a range of values from 2.5 to 7 $\mu g/m^3$.

F. Methodology for estimating ambient concentrations of PM from diesel-fueled engine emissions

The following is a summary of an updated method for estimating ambient diesel PM (DPM) concentrations from ambient NO_x concentrations. A full discussion of the methodology can be found in the Appendix 3. It consists of a simple variation of a receptor model, which uses measurements of ambient chemical concentrations to infer source contributions, known as the tracer species method. A basic assumption in this method is that the ambient concentration of a tracer species, C, may be used alone to infer the ambient concentration of a pollutant from a specific source, S:

$$S = \alpha C,$$

where α is a scale factor that is independent of location. In the estimation of DPM, we take C to be the ambient concentration of NO_x and S to be the ambient concentration of DPM less than 2.5 µm (DPM2.5). The factor α relates the concentration of PM produced by diesel-fueled engine emissions to the concentration of NO_x produced by all sources.

The estimates of the ratio DPM/NO_x from the emission inventory (EI)-population weighted and source apportionment (SA) studies compare very well: EI 0.023 (0.003 or 0.006) and SA South Coast Air Basin 0.026 (0.006) and San Joaquin Valley 0.027 (0.008). This agreement between EI and SA estimates for α, coupled with the uncertainty intervals, motivates the use of a single scaling factor for the whole state of

California to estimate annual average concentrations of DPM from annual average measurements of NO_x. We take the EI values for the average and standard deviations for high and low-NO_x emission counties as best estimates for a population weighted value of DPM/ NO_x: α = 0.023 (0.003 high NO_x counties or 0.006 low-NO_x counties). The value of α gives a population weighted estimate of DPM/ NO_x for all locations in California; the standard deviation values indicate the uncertainty in this choice of α for a given county (based on population).

Based on the agreement between source apportionment and emissions inventory estimates of the scaling factor α, the ratio DPM/total NO_x, we propose the use of a single value of α for estimating the population-weighted annual average ambient DPM concentration for California from NO_x concentrations.

The proposed method to estimate ambient DPM concentrations has distinct advantages over the previous PM10 method (CARB, 1998) as well as several important limitations. The primary strengths of the method include the strong relation of DPM to (total) NO_x, simple application, estimates of uncertainty intervals, and ability to capture sub-county variations in DPM concentrations. In addition to these strengths, the approach is tied directly to the ARB emission inventory, and links bottom-up EI estimates with top-down SA estimates. Several limitations and caveats also bear on applications of the method. The limitations include all assumptions sufficient for application of EI estimates to ambient air, such as well-mixed air parcels (county scale), proportional removal rates for NO_x and DPM, proportionally uniform emission rates for all NO_x and DPM sources, etc. Verification of these assumptions is in general not possible; instead, agreement between EI and SA estimates is taken as best available evidence. The uncertainty intervals produced by the estimation method are based on variations between similar (low- or high- NO_x) counties and reflect differences in relative emission sources (primarily diesel vs. non-diesel). As such, the uncertainty describes the confidence in α to accurately describe either low- or high- NO_x counties. Further work is needed in strengthening the understanding of the contribution of various emission sources to ambient concentrations of both gases and particles. In this respect, source apportionment work that utilizes organic marker species is the best available approach; ideally, highly time-resolved studies would allow better characterization and support for single species scaling estimates, such as the NO_x-scaling method. The following is a discussion of the NO_x data used in this methodology.

Nitrogen Oxides Air Quality Data

Nitrogen dioxide (NO_2) and nitric oxide (NO) are products of all types of combustion. NO reacts with hydrocarbons in the presence of sunlight to form NO_2. Routine ambient air nitrogen oxides are monitored continuously at more than 114 sites in California using federally approved chemiluminescence methods. The data for each monitoring site are reported as 1-hour average concentrations. Statewide estimates of annual average nitrogen oxides concentrations were calculated using data from routine and special monitoring programs, which are briefly described below.

- Continuous hourly measurements of nitrogen oxides data from the 12 Children's Health Study (CHS) air quality monitoring network located in the southern California. NO_2 was determined hourly from EPA-approved chemiluminescent

instruments measuring NO_x and NO.

- Continuous hourly measurements of nitrogen oxides data from the California Regional PM10/PM2.5 Air Quality Study (CRPAQS); measurements were made at a time resolution of 5 or 10 minutes using a gas chromatograph and luminol chemiluminescence detector.

At rural sites, in the absence of nitrogen oxides measurements, the best estimates were obtained using ammonium nitrate data from the Interagency Monitoring of Protected Visual Environments (IMPROVE) program monitoring sites. IMPROVE sites are located in federally protected Class 1 areas and are outside of urban areas. The IMPROVE sampler is programmed to collect two 24-hour duration samples per week. In this data analysis, the mass associated with ammonium nitrate can be estimated by multiplying the nitrate values by the ratio of the molecular weight of ammonium nitrate (80) to the molecular weight of nitrate (62), a factor of 1.29.

From previous data analysis work (Motallebi 2006), a quantitative relationship between precursor emissions and secondary ammonium nitrate was developed. To estimate the conversion of NO_x to PM nitrate, it was suggested that the fraction of NO_x emissions converted to nitrate ranged from 30 to 50 percent. For example, this could indicate that each gram of emitted NO_x produces approximately 0.30 - 0.50 grams of secondary PM (i.e., PM-Nitrate). In this analysis, a mid-range of 40 percent was used to convert ammonium nitrate to NO_x at IMPROVE monitoring sites.

The additional NO_2 data, based on PM nitrate, further improve the spatial coverage of the NO_x monitoring network.

G. Methodology for evaluating risk to small populations exposed to PM emissions from specific sources

Health impacts from PM exposure are commonly estimated at the statewide or a similarly large geographic scale because these estimates are based on epidemiologic studies that relied on single ambient air monitoring stations to represent regional exposures to the pollutant, and incidence rates are calculated at the county level. Our interest is in refining and applying such estimation techniques to finer scales, for small populations being affected by small changes in pollutant concentrations that would result from a single or few sources of emissions.

Below is a summary of two methodologies that are proposed for estimating health impacts associated with exposures to PM resulting from specific sources in a limited geographical area. The discussion is divided into two sections based on available information on the pollutant concentration: a) modeled concentrations and b) emissions data. Examples using ports and goods movement are shown to clarify the discussion.

G.1. Methodology based on modeled concentrations

In the first scenario, an air dispersion model is used to estimate ambient concentrations of PM in a limited geographic area affected by emissions from a specific source or group of sources. Examples would be locomotive emissions at a rail yard or all sources

of diesel (trucks, locomotives, ships) at a California port or harbor. In this scenario, the annual average ambient diesel PM concentration would be estimated by grid cells using a model such as U.S. EPA ISCST3. For each grid cell, the premature death could be estimated based on a C-R function, the population in that grid cell, and the baseline countywide incidence rates. The total impacts for the affected population in the modeling domain would then be obtained by summing the results from each grid cell. In the results section, an example on the Ports of Los Angeles and Long Beach is discussed.

G.2 Methodology based on emissions data only

When it is not feasible to model PM concentrations, emissions can be used to estimate health impacts as an alternative methodology. For example, to estimate health impacts associated with goods movement activities in California, an emissions inventory approach was used in all regions outside of the Ports of Los Angeles and Long Beach, as shown below. Details for this methodology can be found in the CARB 2006 report.

1. Use ARB's estimated county-specific PM2.5 concentrations attributed to diesel sources in year 2000 (CARB 1998).
2. Calculate the premature deaths for the base year 2000 by applying a C-R function to the exposed population for a county.
3. Associate the health impacts with the total diesel PM emission inventory for that county in the base year 2000 to determine the number of tons emitted per annual death. This is called the "tons-per-death" factor for the county.
4. Apply the tons-per-death factor to the diesel PM emission inventory for a single source to estimate the average annual deaths associated only with exposure to these emissions, adjusting for population growth between the year of interest and the base year 2000. Note that the diesel PM emissions from the single source may be small compared to the county's emission inventory used in step 3 above.

III. Results

A. General relationship (relative risk) for use in California

From the procedures described in Section II.D, the central estimate of the relative risk of premature death is 10 percent per 10 $\mu g/m^3$ increase in PM2.5 exposures, with 3 percent to 20 percent confidence interval. The central estimate is the median of the twelve experts' medians (adjusted for the causality likelihood in cases where the expert did not incorporate the likelihood directly into his distribution) from U.S. EPA's expert elicitation, while the lower and upper bounds are the medians of the experts' 5th percentiles and 95th percentiles, respectively. These three values represent our proposed credible range (or uncertainty interval) for the PM2.5-mortality C-R function.

After our credible range was developed, the results from the European Expert Elicitation on the likely relationship between long-term PM2.5 exposures and premature death in the United States were published (Cooke et al. 2007). The median of the six selected European experts' medians is also 10 percent per 10 $\mu g/m^3$ change in PM2.5 exposures, confirming the reasonableness of our central estimate of 10 percent.

A.1 Results of Sensitivity Analyses

To demonstrate the robustness of the relative risk described above, we performed sensitivity analyses using alternative approaches described in Section II.E (Table 3). For each of the alternative scenarios considered, Table 3 presents results in terms of percent change in premature death per 10 μg/m^3 change in PM2.5 exposures, with low indicating 5th percentile and high indicating 95th percentile. For reference, our proposed credible range of the PM2.5-mortality C-R function is listed in the first row. These results showed that regardless of the method chosen, the mean factor relating PM2.5 exposure to premature death lies between 9.5 percent and 12 percent, which brackets our estimate of 10 percent.

Table 3: Percent change in mortality risk per 10 μg/m^3 increase in PM2.5 exposure

Scenario	*Low*	*Mean*	*High*
Proposed credible range	3%	10%	20%
1. 3 studies, equal weight	2%	12%	26%
2. 3 studies, inverse-variance weight	4%	11%	19%
3. 3 studies, random effects pooling	3%	11%	19%
4. 2 studies, random effects pooling	3%	10%	20%
5. twelve experts, random effects pooling	0%*	10%	21%
6. 10 experts, random effects pooling	0%*	9.5%	19%

*Whenever the lowest value in an expert's distribution includes zero, a pooled result (including this expert) can have zero as a lower bound.

B. Results on premature deaths associated with exposures to ambient PM

In this section, we present the results of estimating premature deaths associated with ambient PM exposures above certain cut-off levels. Tables 4a and 4b show the number of premature deaths using a 10 percent relative risk associated with PM2.5 exposures above 7 μg/m^3 and 2.5 μg/m^3, respectively. The level of 7 μg/m^3 represents a reasonable cut-off level based on empirical epidemiologic evidence; and 2.5 μg/m^3 represents the PM2.5 background, the concentration of PM2.5 in the absence of any PM from anthropogenic sources. For this analysis, PM2.5 monitoring data from years 2004 through 2006 were used to represent current ambient PM levels. The population data from the 2000 Census were grown to each corresponding year in BenMAP. As explained in Section II.E.1 above, the results are averages of annual impacts. All results greater than 100 have been rounded to two significant figures. As such, the totals may not add up. The significance of the variation in the results shown in Tables 4a-4d is discussed in Section IV.

Table 4a: Annual premature deaths associated with ambient PM2.5 levels above 7 µg/m^3 *

Air Basin	*Low*	*Mean*	*High*
Great Basin Valleys	<1	1	1
Lake County	<1	1	2
Lake Tahoe	<1	<1	<1
Mojave Desert	37	120	220
Mountain Counties	18	59	110
North Central Coast	5	15	28
North Coast	7	25	45
Northeast Plateau	1	4	8
Sacramento Valley	260	850	1,500
Salton Sea	30	100	180
San Diego County	260	870	1,600
San Francisco Bay	530	1,800	3,200
San Joaquin Valley	610	2,000	3,500
South Central Coast	76	250	460
South Coast	2,500	8,100	14,000
Statewide Total	**4,300**	**14,000**	**25,000**

*Totals do not add up due to rounding; air quality data from years 2004 to 2006.

Table 4b: Annual premature deaths associated with ambient PM2.5 levels above 2.5 µg/m^3 *

Air Basin	*Low*	*Mean*	*High*
Great Basin Valleys	3	8	15
Lake County	6	21	38
Lake Tahoe	1	3	6
Mojave Desert	100	330	580
Mountain Counties	62	210	370
North Central Coast	57	190	340
North Coast	40	130	230
Northeast Plateau	4	15	26
Sacramento Valley	510	1,700	2,900
Salton Sea	89	300	520
San Diego County	510	1,700	3,000
San Francisco Bay	1,100	3,700	6,600
San Joaquin Valley	900	2,900	5,000
South Central Coast	200	670	1,200
South Coast	3,600	12,000	20,000
Statewide Total	**7,200**	**24,000**	**41,000**

*Totals do not add up due to rounding; air quality data from years 2004 to 2006.

C. Results on premature deaths avoided by strategies designed to attain ambient air quality standards

In addition to examining the mortality impacts associated with exposures above certain PM2.5 levels, we also assessed the health benefits of attaining the established ambient air quality standards. Tables 4c and 4d presents the annual premature deaths that would be avoided if PM2.5 levels from the years 2004 through 2006 were reduced to attain the national standard of 15 μg/m^3 and the State standard of 12 μg/m^3. For this calculation the cut off level was set at 2.5 μg/m^3. Interestingly, setting the cut off level to 7 μg/m^3 did not alter the results significantly. The number of premature deaths avoided by attaining the national standard decreased from 5,514 to 5,506 (unrounded); for the State standard, it decreased from 9,300 to 9,200. Details on the methodology used in calculating these estimates are provided in Appendix 2.

Table 4c: Annual premature deaths avoided by attainment of the national annual PM2.5 standard of 15 μg/m^3 *

Air Basin	*Low*	*Mean*	*High*
Great Basin Valleys	<1	<1	<1
Lake County	<1	<1	<1
Lake Tahoe	<1	<1	<1
Mojave Desert	8	27	49
Mountain Counties	2	5	10
North Central Coast	<1	<1	<1
North Coast	<1	<1	<1
Northeast Plateau	<1	<1	<1
Sacramento Valley	4	15	26
Salton Sea	<1	1	1
San Diego County	<1	2	3
San Francisco Bay	1	5	9
San Joaquin Valley	310	1,000	1,900
South Central Coast	1	4	8
South Coast	1,300	4,400	7,900
Statewide Total	**1,700**	**5,500**	**9,900**

*Totals do not add up due to rounding; air quality data from years 2004 to 2006.

Table 4d: Annual premature deaths avoided by attainment of the State annual PM2.5 standard of 12 µg/m^3 *

Air Basin	*Low*	*Mean*	*High*
Great Basin Valleys	<1	<1	<1
Lake County	<1	<1	1
Lake Tahoe	<1	<1	<1
Mojave Desert	12	41	74
Mountain Counties	8	27	48
North Central Coast	1	4	7
North Coast	1	5	8
Northeast Plateau	<1	<1	<1
Sacramento Valley	130	420	760
Salton Sea	16	55	100
San Diego County	94	320	570
San Francisco Bay	210	710	1,300
San Joaquin Valley	460	1,500	2,700
South Central Coast	14	46	83
South Coast	1,900	6,200	11,000
Statewide Total	**2,800**	**9,300**	**17,000**

*Totals do not add up due to rounding; air quality data from years 2004 to 2006.

D. Results on premature deaths associated with exposures to diesel PM exposures

Table 5 lists the estimated premature deaths associated with exposure to diesel PM by air basin. The estimates reflect the central estimate of the relative risk of premature death of 10 percent per 10 µg/m^3 increase in PM2.5 exposures, with 3 percent to 20 percent confidence interval. The concentration of ambient diesel PM concentrations were calculated using the ambient NO_x concentrations. A full discussion of the methodology for estimating diesel PM from NO_x concentrations can be found in the Appendix 3.

Table 5: Annual premature deaths associated with exposures to estimated primary diesel PM*

Air Basin	*Low*	*Mean*	*High*
Great Basin Valleys	<1	<1	1
Lake County	3	9	17
Lake Tahoe	<1	1	2
Mojave Desert	19	66	120
Mountain Counties	8	26	48
North Central Coast	7	22	40
North Coast	4	14	26
Northeast Plateau	<1	<1	<1
Sacramento Valley	55	190	340
Salton Sea	12	40	72
San Diego County	81	270	490
San Francisco Bay	190	640	1,200
San Joaquin Valley	84	280	510
South Central Coast	22	76	140
South Coast	690	2,300	4,100
Statewide Total	**1,200**	**3,900**	**7,100**

*Year 2000, based on the new PM2.5-mortality relative risk of 10 percent per 10 $\mu g/m^3$ increase in PM2.5 exposures. Totals do not add up due to rounding.

E. Results on premature deaths associated with exposures to specific sources

In this section, results are presented based on applications of two methodologies discussed in sections G.1 and G.2.

Ports of Los Angeles and Long Beach. We applied the methodology using modeled concentrations of diesel PM2.5 to assess the mortality effects (described in section G.1) to the area surrounding the Ports of Los Angeles and Long Beach. Using the new PM2.5-mortality function of 10 percent per 10 $\mu g/m^3$ change in PM2.5 exposures, staff estimated that based on modeled diesel PM concentrations for year 2002, the annual premature deaths associated with the ports' emissions are around 120, with uncertainty interval 36 to 310 deaths. The population data from the 2000 Census was grown to estimate the year 2002 populations affected. Details on the modeling methodology used can be found in the CARB 2006 report.

Goods Movement in California. We also used the emissions-based methodology (described in section G.2) to estimate the total mortality impacts associated with PM2.5 generated from all ports and goods movement activities in California. Details on the emissions related to goods movement are in the CARB 2006 report. Using this methodology, staff estimates that annually 3,700 premature deaths can be associated with PM2.5 exposure from goods movement activity statewide. Also noteworthy is that 2,000 premature deaths are associated with exposures to primary diesel PM from goods movement activities, which is slightly more than one-half the total estimated diesel PM impact (from all sources) shown in Table 5.

Table 6: Annual premature deaths associated with PM2.5 from Goods Movement activities[1]

Pollutant	*Low*	*Mean*	*High*
Primary Diesel PM	600	2,000	3,500
Secondary Diesel PM (Nitrates)	480	1,600	2,800
Secondary Diesel PM (Organic Aerosols)	15	49	85
Other Primary PM2.5[2]	12	39	68
Statewide Total[3]	***1,100***	***3,700***	***6,500***

[1]For the year 2005, these estimates do not include the contributions from particle sulfate formed from SO_X emissions, which is being addressed with several ongoing emissions, measurement, and modeling studies. Results listed are based on the previous emission inventories used in the Goods Movement Emission Reduction Plan in April of 2006 but with the new PM2.5-mortality relationship of 10 percent per 10 µg/m^3 increase in PM2.5 exposures; these values may change if emissions inventories are updated.
[2]PM2.5 includes tire wear, brake wear, and particles from boilers, which are not covered under primary diesel PM.
[3]Totals do not add up due to rounding.

IV. Discussion

By evaluating the recent epidemiologic data and the results of the U.S. EPA's expert elicitation, we were able to systematically develop a new range for the relationship between long-term exposures to PM2.5 and the risk for premature death.

Up to now, CARB staff has calculated mortality impacts associated with PM2.5 exposures based on the C-R relationship from the American Cancer Society study (Krewski et al. 2003, Pope et al. 2002). Several recently published studies prompted CARB to consider updating the C-R function as well as other aspects of the methodology for quantifying mortality impacts. In this report, all relevant literature on PM2.5 mortality was reviewed and evaluated, and a new C-R function of 10 percent per 10 µg/m^3 change in PM2.5 exposure was developed (with uncertainty interval from 3 percent to 20 percent). Although the interpretation of the recent literature mostly favors a no-threshold model, staff discussed several possible cut-off levels and presented a range of results. Because of uncertainties in the cut-off concentration below which we expect no PM2.5-related adverse health impacts, staff recommends using a range of cut-off values from 2.5 to 7 µg/m^3. Depending on the cut-off level chosen, as shown in Tables 4a and 4b above, exposures to 2004-06 PM2.5 can be associated with about 14,000 to 24,000 premature deaths statewide annually, with uncertainty ranging from 4,300 to 41,000 deaths.

The methodology for estimating the premature deaths avoided by attaining the ambient PM2.5 annual standards has also been updated. With the new C-R function applied to the updated methodology, about 5,500 deaths (uncertainty: 1,700 to 9,900) are avoided annually if the current PM levels (years 2004 through 2006) are reduced statewide to attain the national standard of 15 µg/m^3. Similarly, about 9,300 deaths (uncertainty: 2,800 to 17,000) would be avoided if the State standard of 12 µg/m^3 is attained

statewide.

Treating diesel PM and ambient PM as equally toxic and using the new PM2.5-mortality function, staff estimate that statewide, public exposures to diesel PM can be associated with about 3,900 deaths, with uncertainty ranging from 1,200 to 7,100.

The PM2.5-mortality concentration-response function we developed can be applied in regional (i.e., by county) assessments of premature deaths associated with PM2.5 exposures, as most epidemiological studies relate death and health data with regional PM measurements that apply to large populations. However, recent advances in exposure classification techniques, as demonstrated by Jerrett et al. (2005) for example, suggest that it is also reasonable to apply the PM2.5-mortality relationship to analyses involving populations of small sizes, as long as uncertainties and limitations are explicitly stated. Staff demonstrated such applications in estimating the mortality impacts associated with PM2.5 emissions related to port activities for the Ports of Los Angeles and Long Beach. Using the new PM2.5-mortality relationship, it is estimated at about 110 premature deaths (uncertainty interval: 36 to 310) are associated with annual PM2.5 exposures to emissions resulting from such activities.

It should be noted that while this report focuses on premature death, additional quantified health impacts include hospital admissions, lost workdays, minor restricted activity days, and a number of other health endpoints (CARB 2006). Still, some other health effects (e.g. asthma exacerbation) cannot be quantified at this time (CARB 2006). Therefore, taken as a whole, the overall health benefits of PM reduction may be under-estimated.

V. Uncertainties and Limitations

There are a number of uncertainties involved in quantitatively estimating the health impacts associated with exposures to outdoor air pollution. Over time, some of these will be reduced as new research is conducted. However, some uncertainty will remain in any estimate. Below, some of the major uncertainties and limitations of the estimated health impacts presented in this report are briefly discussed.

<u>Concentration-Response Function</u>

A primary uncertainty is the choice of the specific studies and the associated concentration-response (C-R) functions used for quantification. Epidemiological studies used in this report have undergone extensive peer review and include sophisticated statistical models that account for the confounding effects of other pollutants, meteorology, and other factors. While there may be questions on whether C-R functions from the epidemiological studies are applicable to California, it should be noted that some of the cities in the ACS cohort are in California. Also, studies have shown that the mortality effects of PM in California are comparable to those found in other locations in the United States (Dominici et al. 2005, Franklin et al. 2007, Jerrett et al. 2005; Pope et al. 2002). The C-R function for PM2.5-related mortality developed in this report was based on a careful review of all relevant scientific literature and a thorough consideration of their strengths and limitations. In addition, it was approved by our advisors and independent peer reviewers.

Many of the studies were conducted in areas having fairly low concentrations of ambient PM, with ranges in PM levels that cover California values. Thus, the extrapolation is within the range of the studies. Finally, the uncertainty in the C-R functions selected is reflected in the lower and upper estimates given in all the health impacts tables, which represent 95 percent confidence intervals.

Baseline Mortality Rate

Mortality baseline rates are entered into the C-R functions in order to calculate the estimates presented in this report, and there is uncertainty in these baseline rates. Often, one must assume a baseline incidence level to be consistent throughout the city or county of interest. In addition, incidence can change over time as lifestyles, income and other factors evolve. For this analysis, we utilized baseline rates that are used by U.S. EPA. Additional information were obtained from Department of Health Services and the Centers for Disease Control and Prevention. It is expected that incidence rates may change over time.

Diesel PM Compared to Ambient PM Relative Toxicity

In this assessment, staff assumed diesel PM is equally toxic as PM2.5. Without definitive evidence to conclude otherwise, this approach may underestimate the true effects of diesel PM exposures on adverse health effects.

Diesel PM Concentrations

In the absence of a direct measurement method, ambient diesel PM concentrations were estimated from ambient NO_x concentrations. These diesel PM estimates depend upon the network of ambient NO_x measurements from ARB monitoring sites. A basic assumption in this method is that the ambient concentration of a tracer species may be used to infer the ambient concentration of diesel PM.

The limitations include all assumptions sufficient for application of emissions inventory estimates to ambient air, such as well-mixed air parcels (county scale), proportional removal rates for NO_x and diesel PM, proportionally uniform emission rates for all NO_x and diesel PM sources. Verification of these assumptions is in general not possible. Instead, agreement between emissions inventory and source apportionment estimates is taken as best available evidence. The uncertainty intervals produced by the estimation method are based on variations between low- NO_x counties and reflect differences in relative emission sources (primarily diesel vs. non-diesel). However, this uncertainty has not been incorporated into estimating the premature deaths associated with diesel PM exposure in this report.

Interpolation

Interpolation is the procedure of predicting the PM2.5 concentration at areas without ambient measurements. Interpolation is necessary when monitoring data do not cover the area of interest completely. The source of error for this analysis stems from measurement error and error associated with having enough monitors to get adequate spatial coverage. When data are abundant, most interpolation techniques give similar results. When data are sparse, however, the assumption made about the underlying variation that has been sampled and the choice of method and its parameters can be

critical if one is to avoid misleading results.

Exposure concentration

There are three methods for estimating the exposure concentration used to estimate PM2.5-related mortality: ambient measurement, modeled concentration and emissions inventory. There are advantages, uncertainties, and limitations with each method.

Concentration is estimated from ambient measurement by interpolating in areas with no measured concentration. The technique used in this report was inverse distance weighted squared. It has the advantage of having a high degree of certainty of the pollutant concentration near the monitoring station. As the distance increases away from the monitoring station, the uncertainty in the interpolated concentration also increases. In areas with high spatial coverage and low variability in concentration, this method gives the most reliable estimate of concentration.

When ambient measurements are not available, modeled concentration estimates of ambient air quality are done using emission inventories and air quality models. The models may be simple box models that track the movement of an air parcel through a region or detailed models that incorporate photochemical reactions and complex terrain. This technique has the advantage of estimating the relative source of PM2.5 compared to other sources. It can, for example, estimate the amount of PM2.5 from ships, trucks, or stationary sources at a particular location. Modeling can also estimate localized concentrations with sharp gradients that would not be feasible to measure with air quality monitors. The downside to modeling is that it is labor intensive and has an uncertainty of about a factor of two. Nonetheless, it is the next best tool when ambient monitoring is not feasible.

The least reliable estimation of health impacts occurs when emissions are used to infer about air quality. As outlined in section II.G.2, this method estimates the health benefits associated with reductions in PM2.5 emissions due to ARB regulatory action. To infer health impacts due to emission reductions, this method applies a “tons of PM2.5 per death” factor to estimate the number of deaths avoided due to reductions in PM2.5. The method may give an overestimate of mortality where sources are far from populated areas. For example, emissions from the Ports of Los Angeles and Long Beach are miles away from populated areas, and would result in an overestimate of mortality. It may also produce an underestimate where the source of PM2.5 is in close proximity to populated sources.

VI. Conclusions

This report was a product of an evaluation of the available published literature on PM mortality. A new relative risk factor of premature death associated with PM2.5 exposures was developed: 10 percent increase in premature death per 10 $\mu g/m^3$ increase in PM2.5 exposures (uncertainty interval: 3 percent to 20 percent). Also, staff proposed to use a range of cut-off levels between 2.5 to 7 $\mu g/m^3$ based on the health effects observed over the range of PM concentrations recorded in case-control studies.

Although the literature mostly favors a no-threshold model, without strong empirical evidence for long-term PM effects between 2.5 and 7 $\mu g/m^3$, we recommend this range for the purpose of assessing the premature deaths associated with long-term exposures to fine PM. The methodologies and results presented in this report have been endorsed by our scientific advisors and have undergone an external peer review process.

VII. References

Abbey, D.E.; Nishino, N.; McDonnell, W.F.; Burchette, R.J.; Knusten, S.F.; Beeson, W.L.; Yang, J.X. Long-Term Inhalable Particles and Other Air Pollutants Related to Mortality in Nonsmokers; *Am. J. Respir. Crit. Care Med.* (1999), *159,* 373-382.

APHEA2. In *Revised Analyses of Time-Series of Air Pollution and Health. Special Report*; Health Effects Institute: Boston, MA, (2003); pp 157–164.

Beelen, R.; Hoek, G.; van den Brandt, P.A.; Goldbohm, R.A.; Fischer, P.: Schouten, L.J.; Jerrett, M.; Hughes, E.; Armstrong, B.; Brunekreef, B. Long-Term Effects of Traffic-Related Air Pollution on Mortality in a Dutch Cohort (NLCS-AIR Study); Environmental Health Perspectives (2008), 116:196-202.

Brunekreef, B.; Hoek, G. A Critique of "Fine Particulate Air Pollution and Total Mortality Among Elderly Californians, 1973-2002" by James E. Enstrom, PhD; *Inhalation Toxicology* (2006), 18:507-508.

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

Burnett, R.T.; Goldberg, M.S. Size-Fractionated Particulate Mass and Daily Mortality in Eight Canadian Cities. In *Revised Analyses of Time-Series of Air Pollution and Health. Special Report*; Health Effects Institute: Boston, MA, (2003); pp 85–90.

CARB 1998. California Air Resources Board, Proposed Identification of Diesel Exhaust as a Toxic Air Contaminant. Appendix III. Part A: Exposure Assessment, available at http://www.arb.ca.gov/toxics/id/summary/diesel_a.pdf. (1998)**.**

CARB 2002. California Air Resources Board and Office of Environmental Health Hazard Assessment. Staff Report: Public Hearing to Consider Amendments to the Ambient Air Quality Standards for Particulate Matter and Sulfates, available at http://www.CARB.ca.gov/research/aaqs/std-rs/pm-final/pm-final.htm. (2002).

CARB 2003a. California Air Resources Board. Staff Report: Proposed Diesel Particulate Matter Control Measure For On-Road Heavy-Duty Residential and Commercial Solid Waste Collection Vehicles, available at
http://www.CARB.ca.gov/regact/dieselswcv/isor3.pdf. (2003)

CARB 2003b. California Air Resources Board. Staff Report: Proposed Airborne Toxic Control Measure For In-Use Diesel-Fueled Transport Refrigeration Units (TRU) And TRU Generator Sets, And Facilities Where TRUs Operate, available at http://www.CARB.ca.gov/regact/trude03/isor.pdf. (2003)

CARB 2003c. California Air Resources Board. Staff Report: Airborne Toxic Control Measure For Stationary Compression-Ignition Engines, available at http://www.CARB.ca.gov/regact/statde/isor.pdf. (2003)

CARB 2004a. California Air Resources Board. Staff Report: Proposed Modifications To The Fleet Rule For Transit Agencies And New Requirements For Transit Fleet Vehicles, available at http://www.CARB.ca.gov/regact/bus04/isor.pdf. (2004)

CARB 2004b. California Air Resources Board. Staff Report: Airborne Toxic Control Measure For Diesel-Fueled Portable Engines, available at http://www.CARB.ca.gov/regact/porteng/isor.pdf. (2004)

CARB 2004c. California Air Resources Board. Staff Report: Proposed Regulatory Amendments Extending the California Standards for Motor Vehicle Diesel Fuel to Diesel Fuel Used in Harborcraft and Intrastate Locomotives, available at http://www.CARB.ca.gov/regact/carblohc/isor.pdf. (2004)

CARB 2005. California Air Resources Board, "Diesel PM Exposure Assessment Study for the Ports of Los Angeles (POLA) and Long Beach (POLB)", Oct 2005, available at http://www.CARB.ca.gov/msprog/offroad/marinevess/documents/100305draftexposrep.pdf. (2005).

CARB 2006. California Air Resources Board, "Quantification of the Health Impacts and Economic Valuation of Air Pollution From Ports and Goods Movement in California." Appendix A in Emission Reduction Plan for Ports and Goods Movement. March 22, (2006), available at http://www.CARB.ca.gov/planning/gmerp/march21plan/appendix_a.pdf

Chen, L.H.; Knutsen, S.F.; Shavlik, D.; Beeson, W.L.; Petersen, Ghamsary, M.; Abbey, D. The Association between Fatal Coronary Heart Disease and Ambient Particulate Air Pollution: Are Females Greater Risk? *Environ. Health Perspect.* (2005), *113,* 1723-1729.

Clancy, L.; P. Goodman, et al.; Effect of air-pollution control on death rates in Dublin, Ireland: an intervention study. *Lancet* (2002), 360(9341): 1210-4.

Cooke, R.M; Wilson, A.M.; Toumisto, J.T.; Morales, O.; Tainio, M.; Evans, J.S. A Probabilistic Characterization of the Relationship between Fine Particulate Matter and Mortality: Elicitation of European Experts. Env. Sci. and Tech (2007) 41:6598-6605.

de Kok, T.; Driece, H.; Hogervorst, J.; Briede, J.. Toxicological assessment of ambient and traffic-related particulate matter: a review of recent studies. Mutation Research (2006), 613:103-122.

Dockery, D.W.; Pope, C.A., III; Xu, X.; Spengler, J.D.; Ware, J.H.; Fay, M.E.; Ferris, B.G.; Speizer, F.A. An Association between Air Pollution and Mortality in Six U.S. Cities; *N. Engl. J. Med.* (1993), *329,* 1753-1759.

Dominici F, McDermott A, Daniels M, Zeger SL, Samet JM. Revised analyses of the National Morbidity, Mortality, and Air Pollution Study: mortality among residents of 90 cities. J Toxicol Environ Health A. (2005);68(13-14):1071-92.

Enstrom, J.E. Fine Particulate Air Pollution and Total Mortality Among Elderly Californians, 1973–2002; *Inhal. Toxicol.* (2005), *17,* 803-816.

Filleul, L.; Rondeau, V.; Vandentorren, S.; Le Moual, N.; Cantagrel, A.; Annesi-Maesano, I.; Charpin, D.; Declercq, C.; Neukirch, F.; Paris, C.; Vervloet, D.; Brochard, P.; Tessier, J.F.; Kauffmann, F.; Baldi, I. Twenty-Five Year Mortality and Air Pollution: Results from the French PAARC Survey; *Occup. Environ. Med.* (2005), *62,* 453-460.

Finkelstein, M.M.; Jerrett, M.; Sears, M.R. Traffic Air Pollution and Mortality Rate Advancements Periods; *Am. J. Epidemiol.* (2004), *160,* 173-177.

Franklin M, Zeka A, Schwartz J. Association between PM(2.5) and all-cause and specific-cause mortality in 27 US communities. J Expo Sci Environ Epidemiol. (2007), 17, 279-287.

Ghio, A.J.; Biological effects of Utah Valley ambient air particles in humans: a review. *Journal of Aerosol Medicine* (2004), 17(2): 157-164.

Goss, C.H.; Newsom, S.A.; Schildcrout, J.S.; Sheppard, L.; Kaufman, J.D. Effect of Ambient Air Pollution on Pulmonary Exacerbations and Lung Function in Cystic Fibrosis; *Am. J. Repir. Crit. Care Med.* (2004), *169,* 816-821.

Hedley, A.J.; C.M. Wong, et al.; Cardiorespiratory and all-cause mortality after restrictions on sulphur content of fuel in Hong Kong: an intervention study. *Lancet* (2002), 360(9346): 1646-1652.

Holmstead, J.; U.S. EPA, Office of Air and Radiation. Letter to Robert O'Keefe, Health Effects Institute. July 11, (2005).

Hoek, G.; Brunekreef, B.; Goldhohm, S.; Fischer, P.; van den Brandt, P.A. Association between Mortality and Indicators of Traffic-Related Air Pollution in the Netherlands: A Cohort Study; *Lancet* (2002), *360,* 1203-1209.

IE, 2004. An expert judgment assessment of the concentration-response relationship between PM2.5 exposure and mortality. Industrial Economics, Incorporated, prepared for the Office of Air Quality Planning and Standards, Research Triangle Park, NC. (2004), available at http://www.epa.gov/ttn/ecas/regdata/Benefits/pmexpert.pdf.

IE, 2006. Expanded Expert Judgment Assessment of the Concentration-Response Relationship Between PM2.5 Exposure and Mortality, Industrial Economics, Incorporated, Final Report, September 21, (2006), available at http://www.epa.gov/ttn/ecas/regdata/Uncertainty/pm_ee_report.pdf.

Jerrett, M.; Burnett, R.T.; Willis, A.; Krewski, D.; Goldberg, M.S.; DeLuca, P.;

Finkelstein, N. Spatial Analysis of the Air Pollution-Mortality Relationship in the Context of Ecologic Confounders; *J. Toxicol. Environ. Health* (2003), *66,* 1735-1777.

Jerrett, M.; Burnett, R.T.; Ma, R.; Pope, C.A., III; Krewski, D.; Newbold, K.B.; Thurston, G.; Shi, Y.; Finkelstein, N.; Calle, E.E.; Thun, M.J. Spatial Analysis of Air Pollution and Mortality in Los Angeles; *Epidemiol.* (2005a), *16,* 727-736.

Jerrett, M.; Finkelstein, M. Geographies of Risk in Studies Linking Chronic Air Pollution Exposure to Health Outcomes. *J. Toxicol. Environ. Health A* (2005b), *68,* 13-14, 1207-1242.

Kappos, A.; Bruckmann, P.; Eikmann, T.; Englert, N.; Heinrich, U.; Hoppe, P.; Koch, E.; Krause, G.; Kreyling, W.; Rauchfuss, K.; Rombout, P.; Schulz-Klemp, V.; Thiel, W.; Wichmann, H.E. Health effects of particles in ambient air. *Int. J. Hyg. Environ. Health* (2004), 207, 399 - 407

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

Laden, F.; Schwartz, J.; Speizer, F.E.; Dockery, D.W. Reduction in Fine Particulate Air Pollution and Mortality: Extended Follow-Up of the Harvard Six Cities Study. *Am. J. Respir. Crit. Care Med.* (2006), *173,* 667-672.

Lipfert, F.W.; Perry, H.M., Jr.; Miller, J.P.; Baty, J.D.; Wyzga, R.E.; Carmody, S.E. The Washington University-EPRI Veterans' Cohort Mortality Study: Preliminary Results; *Inhal. Toxicol.* (2000), *12,* 41-73.

Lipfert, F.W.; Perry, H.M., Jr.; Miller, J.P.; Baty, J.D.; Wyzga, R.E.; Carmody, S.E. Air Pollution, Blood Pressure, and Their Long-Term Associations with Mortality; *Inhal. Toxicol.* (2003), *15,* 493-512.

Lipfert, F.W.; Wyzga, R.E.; Baty, J.D.; Miller, J.P. Traffic Density as a Surrogate Measure of Environmental Exposures in Studies of Air Pollution Health Effects: Long-Term Mortality in a Cohort of US Veterans; *Atmos. Environ.* (2006a), *40,* 154-169.

Lipfert, F.W.; Baty, J.D.; Miller, J.P.; Wyzga, R.E. PM2.5 Constituents and Related Air Quality Variables as Predictors of Survival in a Cohort of US Veterans; *Inhal. Toxicol..* (2006b), 18*,* 645-657.

Mallick, R.; K. Fung, et al.; Adjusting for measurement error in the Cox Proportional Hazards Regression Model. *Journal of Cancer Epidemiology and Prevention* (2002)*,* 7(4): 155-164.

McDonnell, W.F.; Nishino-Ishikawa, N.; Petersen, F.F.; Chen, L.H.; Abbey, D.E. Relationships of Mortality with the Fine and Coarse Fraction of Long-Term Ambient PM10 Concentrations in Nonsmokers; *J. Expo. Anal. Environ. Epidemiol.* (2000), *10,* 427-436.

Miller, K.A.; Siscovick, D.S.; Sheppard, L.; Shepherd, K.; Anderson, G.; Kaufman, J.D. Air Pollution and Cardiovascular Disease Events in the Women's Health Initiative Observational (WHI-OS) Study. *Circulation* (2004), 109, e71. (Abstract form the American Heart Association Conference on Cardiovascular Disease Epidemiology and Prevention. Full report currently in review).

Miller, K.A.; Siscovick, D.S.; Sheppard, L.; Shepherd, K.; Anderson, G.; Kaufman, J.D. Long-term Exposure to Air Pollution and Incidence of Cardiovascular Events in Women. *N England J Med* (2007), 365:5, 447-458.

Morgan, M.G., M. Henrion; Uncertainty: a guide to dealing with uncertainty in quantitative risk and policy analysis. (New York: Cambridge University Press). (1990).

Motallebi, N. Conversion Factors for Secondary Formation of PM-Nitrate from NO_x Emissions. An internal report in Research Division, California Air Resources Board, (2006)

Motallebi N., Taylor C.A., and Croes B.E. Particulate matter in California: Part 2 - Spatial, temporal, and compositional patterns of PM2.5, PM10-2.5, and PM10. *J. Air & Waste Manag. Assoc.* (2003), 53 (12), 1517-1530.

Nafstad, P.; Håheim, L.L.; Wisløff, T.; Gram, F.; Oftedal, B.; Holme, I.; Hjermann, I.; Leren, P. Urban Air Pollution and Mortality in a Cohort of Norwegian Men; *Environ. Health Perspect.* (2004), *112,* 610-615.

National Research Council (NRC). Estimating the public health benefits of proposed air pollution regulations. The National Academies Press: Washington, D.C. (2002).

Ostro B.; Browdwin, R.; Green, S.; Feng, W.; Lipsett, M. Fine Particulate Air Pollution and Mortality in Nine California Counties: Results from CALFINE. Env Health Persp (2006) V114:29-33.

Pope, C.A., III; Thun, M.J.; Namboodiri, M.M.; Dockery, D.W.; Evans, J.S.; Speizer, F.E.; Heath, J.C.W. Particulate Air Pollution as a Predictor of Mortality in a Prospective Study of U.S. Adults; *Am. J. Respir. Crit. Care. Med.* (1995), *151,* 669-674.

Pope, C.A.; III; Burnett, R.T.; Thun, M.J.; Calle, E.E.; Krewski, D.; Ito, K.; Thurston, G.D. Lung Cancer, Cardiopulmonary Mortality, and Long-Term Exposure to Fine Particulate Air Pollution; *J. Am. Med. Assoc.* (2002), *287,* 1132-1141.

Pope, C.A.; III; Burnett, R.T.; Thurston, G.D.; Thun, M.J.; Calle, E.E.; Krewski, D.; Godleski, J.J. Cardiovascular Mortality and Long-Term Exposure to Particulate Air Pollution: Epidemiological Evidence of General Pathophysiological Pathways of Disease; *Circulation* (2004), *109,* 71-77.

Pope, C.A.; Respiratory disease associated with community air pollution and a steel

mill, Utah Valley. *American Journal of Public Health* (1989), 79(5): 623-628.

Pope, C.A.; Respiratory hospital admissions associated with PM10 pollution in Utah, Salt Lake, and Cache Valleys. *Archives of Environmental Health* (1991), 46(2): 90-97.

Pope, C. A., M. J. Thun, et al.; Particulate air-pollution as a predictor of mortality in a prospective-study of U.S. adults. *American Journal of Respiratory and Critical Care Medicine* (1995), 151(3): 669-674.

Pope, C.A.; Particulate pollution and health: a review of the Utah valley experience. *Journal of Exposure Analysis and Environmental Epidemiology* (1996), 6(1): 23-34.

Pope, C.A,; III; Dockery, D.W.; Health Effects of Fine Particulate Air Pollution: Lines that Connect; J. Air & Waste Manage. Assoc. (2006) 56:709-742.

Roman, H.A.; Walker, K.D.; Walsh, T.L.; Conner, L.; Richmond, H.M.; Hubbell, B.J.; Kinney, P.L. Expert Judgment Assessment of the Mortality Impact of Changes in Ambient Fine Particulate Matter in the U.S. Env Sci and Tech. Published on Web 02/27/2008.

Samet, J. M.; F. Dominici, et al.; The National Morbidity, Mortality, and Air Pollution Study Part I: Methods and Methodologic Issues. Research Report 94, Health Effects Institute, Boston, MA. (2000a).

Samet, J. M.; S. L. Zeger, et al.; The National Morbidity, Mortality, and Air Pollution Study Part II: Morbidity and Mortality from Air Pollution in the United States. Research Report 94, Health Effects Institute, Boston, MA. (2000b).

Schwartz J.; Dochery, D.; Neas, L.M. Is Daily Mortality Associated Specifically with Fine Particles? Air and Waste Manage Assoc (1996) 46:927-939.

Schwartz, J.; Laden, F.; Zanobetti. The concentration-response relation between PM2.5 and daily deaths. Env Health Persp (2002) 110, 10: 1025-1029.

Sun, Q., A. Wang, et al.; Long-term air pollution exposure and acceleration of atherosclerosis and vascular inflammation in an animal model. *Journal of the American Medical Association* (2005), 294(23): 3003-3010.

U.S. Environmental Protection Agency. (2004). Final regulatory impact analysis: control of emissions from non-road diesel engines. Office of Transportation and Air Quality. EPA-420-R-04-007. (2004).

U.S. Environmental Protection Agency, Science Advisory Board. (2004). Advisory on plans for health effects analysis in the analytical plan for EPA's second prospective analysis – benefits and costs of the Clean Air Act, 1990-2020. Advisory by the Health Effects Subcommittee of the Advisory Council on Clean Air Compliance Analysis. EPA-SAB-COUNCIL-ADV-04-002.

U.S. Environmental Protection Agency. (2005) Regulatory impact analysis for the final Clean Air Interstate Rule. Office of Air and Radiation. EPA-452/R-05-002. (2005).

Walker, K.D.; Industrial Economics, Inc. Memorandum to Jim Neumann, Henry Roman, and Tyra Gettleman, Industrial Economics, Inc. Appropriate Number of Experts for the Particulate Matter Expert Judgment Project. November 11, (2004).

Willis, A.; M. Jerrett, et al.; The association between sulfate air pollution and mortality at the county scale: an exploration of the impact of scale on a long-term exposure study. *Journal of Toxicology and Environmental Health, Part A* (2003) 66(16-19): 1605-1624.

Woodruff, T.J.; Grillo, J.; Schoendorf, K.C. The Relationship between Selected Causes of Postneonatal Infant Mortality and Particulate Air Pollution in the United States; *Environ. Health Perspect.* (1997), *105,* 608-612.

Woodruff, T.J.; Parker, J.D.; Schoendorf, K.C. Fine Particulate Matter (PM2.5) Air Pollution and Selected Causes of Postneonatal Infant Mortality in California. *Environ. Health Perspect.* (2006), *114,* 786-790.

This page is intentionally blank.

APPENDICES

This page is intentionally blank.

Appendix 1 (PM2.5 Exposures)

Below are estimated basin-specific PM 2.5 population-weighted concentrations for years 2004 to 2006 used in this report.

Air Basin	Census 2000 Population	PM2.5 ($\mu g/m^3$) Year 2004	PM2.5 ($\mu g/m^3$) Year 2005	PM2.5 ($\mu g/m^3$) Year 2006
Great Basin Valleys	32,006	6.18	6.69	3.44
Lake County	58,309	4.96	5.17	5.63
Lake Tahoe	46,200	4.31	3.55	3.63
Mojave Desert	816,742	9.16	8.80	8.50
Mountain Counties	408,039	7.60	7.41	8.39
North Central Coast	710,598	7.00	7.12	7.18
North Coast	310,061	7.11	6.98	7.49
Northeast Plateau	87,578	4.91	4.71	5.25
Sacramento Valley	2,334,277	11.41	10.84	11.82
Salton Sea	465,886	9.69	9.55	8.78
San Diego County	2,813,833	12.61	10.98	11.06
San Francisco Bay	6,605,921	11.51	10.70	10.69
San Joaquin Valley	3,189,385	16.32	16.48	16.74
South Central Coast	1,400,455	10.09	9.57	9.23
South Coast	14,592,351	17.57	16.09	14.87
Statewide	***33,871,641***	**14.34**	**13.36**	**12.91**

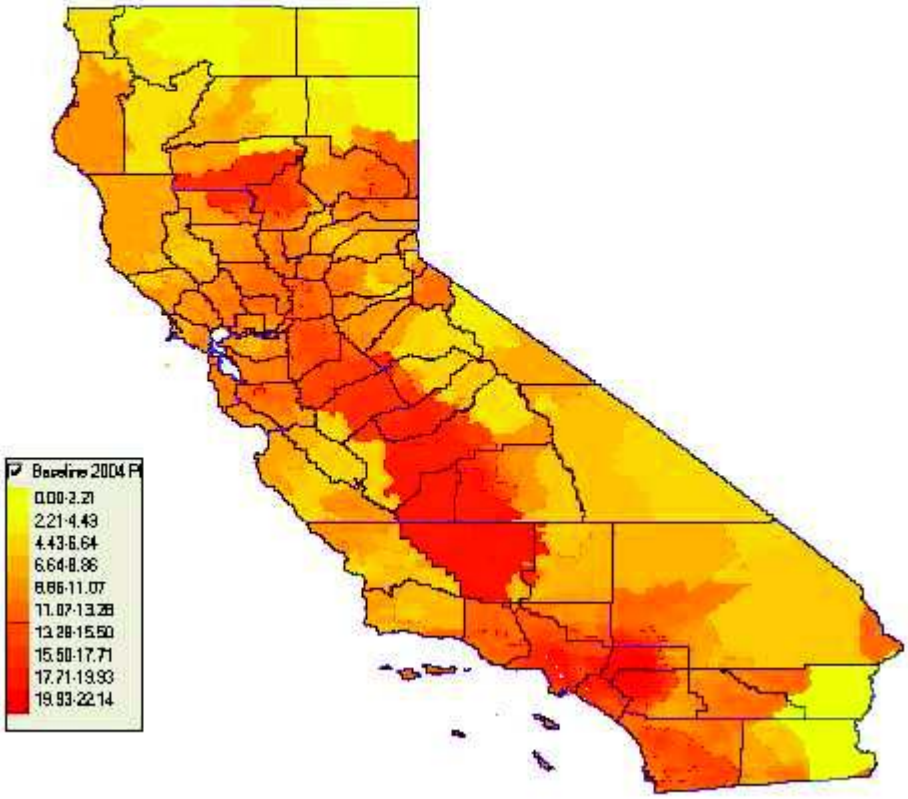
PM2.5 Air Quality Monitoring Program in California

California's air quality monitoring program provides information used for determining which areas violate standards, characterizing the sources that contribute to pollution, determining background concentrations, assessing pollution transport, and supporting health studies and other research. To assess the nature and extent of the PM2.5 problem in California, ARB and air districts have significantly expanded the PM2.5 monitoring program since late 1998. The PM2.5 mass data used in this analysis have been derived from a variety of routine and special monitoring programs and databases. We analyzed the following ambient air quality data:

- 2004-2006 PM2.5 mass from the Federal Reference Method (FRM) monitors. California's PM2.5 monitoring network now includes 90 FRM monitoring sites. The FRM sites collect 24-hour mass data using federally approved methods, which means they satisfy specific federal regulatory requirements.
- 2004-2006 PM2.5 mass data from the Interagency Monitoring of Protected Visual Environments (IMPROVE) program. Since 1985, this program implemented an extensive long term monitoring program to establish the current visibility conditions, track changes in visibility and determine causal mechanism for the visibility impairment in the National Parks and Wilderness Areas. The IMPROVE sampler is programmed to collect two 24-hour duration samples per week.

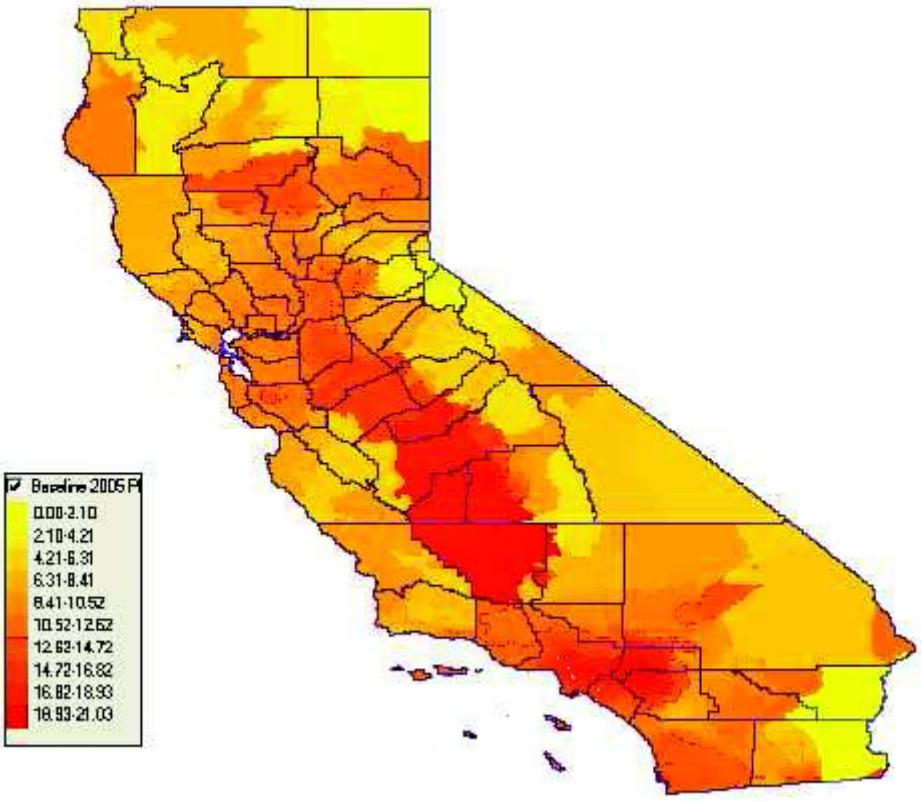
2004 Particulate Matter

Inverse Distance Squared Interpolation



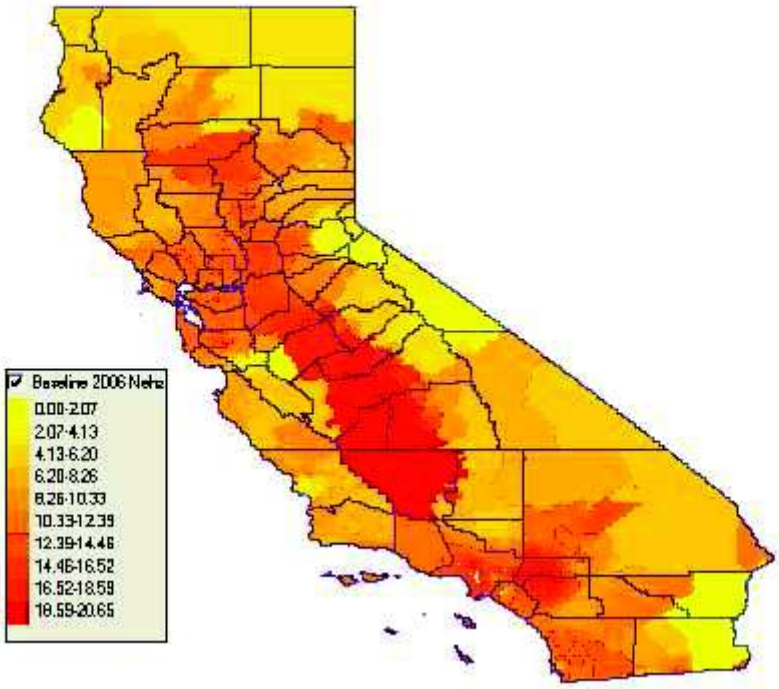
2005 Particulate Matter

Inverse Distance Squared Interpolation



2006 Particulate Matter

Inverse Distance Squared Interpolation



This page is intentionally blank.

Appendix 2 (Methodology for Estimating Health Impacts Avoided by Strategies Designed to Attain the Standards)

In addition to examining the mortality impacts associated with exposures above certain PM2.5 levels, it is important to assess the health benefits of attaining the established ambient air quality standards. When evaluating the impacts associated with implementing strategies designed to attain an ambient air quality standard, we project a future scenario when the highest observed PM measurements are below the established standard – bringing the air basin into attainment of the standard. In this scenario, measurements at all sites within each air basin are also lower compared to current levels; hence their values are "rolled back" to reflect the attainment scenario. As shown in the section labeled "Justification for rollback" below, PM2.5 measurements within several air basins have declined at fairly consistent rates over time, justifying the assumption of a constant rate of reduction within each basin. Details on each step of this "rollback" methodology used to estimate the health impacts avoided by implementing strategies to attain the standards follow.

STEP 1: Obtain PM concentrations for all sites in California

The observed PM2.5 concentrations are obtained for years 2004, 2005, and 2006. In addition to the routine monitoring network, data from the IMPROVE (Interagency Monitoring for Protected Visual Environments) are included in the analysis. See Appendix 1 for a description of these special monitoring data. Annual averages of quarterly means are calculated for each site for consistency with the national and state definition of the PM standard attainment designations.

Consistent with the proportional roll-back procedure applied in the ozone standard staff report (CARB 2004) and published in JAWMA (Ostro et al. 2006), the PM annual averages of quarterly averages are rolled into attainment of a standard as follows.

Denote:

`Current PM`	= current annual PM value
`Basin Max`	= highest value in each basin during 2003-2005
`Background`	= background PM2.5 concentration of 2.5 $\mu g/m^3$
`Standard`	= 15 $\mu g/m^3$ for the federal, 12 $\mu g/m^3$ for the state
`Attainment PM`	= rolled-back PM value in the "attainment" scenario

First, the rollback factor for each basin was calculated as follows:

```
if Basin Max > Standard then
```

$$\text{Rollback Factor} = \frac{\text{Standard - Background}}{\text{Basin Max - Background}}$$

```
else

   Rollback Factor = 1
```

That is, for each air basin, we assumed that only the portion of the PM2.5 average above background will decrease as progress toward attainment of a standard takes place. Thus, for each air basin, the rollback factor represents the percentage reduction needed to bring the basin high towards attainment of a standard.

Next, for all sites within the basin, the portion of the current PM annual average above background was shrunk by the rollback factor, as follows:

```
if Current PM > Background then

   Attainment PM = Background + (Rollback Factor)×(Current PM - Background)
else

   Attainment PM = Current PM
```

The assumption of applying a basin-specific rollback factor to all sites within each basin is justified by the investigation detailed below. Further, it is consistent with air quality plans which are aimed at attaining an appropriate air quality standard by designing programs that would bring down ambient measurements at the high site and at the same time reduce levels at other sites within each basin.

STEP 2: Estimate PM concentration per census tract

The concentration per census tract is estimated using the ambient annual average PM2.5 concentrations measured at monitoring sites. This step is done with BenMAP[9], a software program developed by the U.S. EPA for estimating and mapping health impacts associated with air pollution. BenMAP interpolates PM concentrations using nearby monitored values with the inverse distance weighted squared method.

The interpolation is confined to a 50-kilometer radius, with the weight assigned to each nearby monitored PM value as the inverse square of the distance from the monitor to the location of interpolation. In some areas of California, there may be no monitoring information within 50 kilometers. In these cases, the concentration that will be assigned will be from the closest monitor, regardless of the distance. The end result is a smooth contour surface of PM values throughout the entire state. The interpolated value is then assigned to each census block center. This step is performed for each of the three years.

The same procedure is applied to obtain observed as well as rolled-back exposures in each tract. This step is performed for each of the three years.

STEP 3: Estimate mortality impact

The concentration-response functions are applied to calculate mortality impacts due to long-term changes in PM exposure, using county-specific baseline incidence rates from the Center for Disease Control[10].

[9] http://www.epa.gov/air/benmap/download.html
[10] http://wonder.cdc.gov/mortSQL.html

For log-linear functions, the health impact is

$\Delta Y = -Y_O [\exp(-\beta * \Delta PM - 1)] * pop$, where

Y_O = baseline mortality rates, which include all-case deaths for the population over age 30. We used the mortality rate for the year 2005 to calculate health impacts for years 2004, 2005 and 2006.

β = beta coefficient derived from the relative risk of epidemiologic study results.

ΔPM = the difference between the current ambient PM concentration and the rolled-back or attainment PM level.

pop = population age 30 or above in each census block, from US Census for each year (2004-2006).

Note that the baseline mortality rate and population are available for various subgroups (age 30-34, 35-44, 45-54, 55-64, 65-74, 75-84, 85+). The health impact is actually calculated for each subgroup at the census track level. After each change in health impacts is calculated for each census track, we sum across the results for an air basin or for the entire state. Heath impacts are calculated for each year; they are then averaged over three years to reduce the influence of any year with unusual meteorology on the overall results.

Justification for Rollback

In the discussion above, the roll-back methodology was based on an assumption of a constant rate of PM2.5 reductions within each basin. The validity of this assumption was investigated through an empirical analysis of historical PM2.5 data using various data sources. We examined the rate of decrease in PM levels in Mountain Counties, South Coast, San Francisco Bay Area, San Joaquin Valley, and Sacramento Valley Air Basins, where there were sufficient data between 2000 and 2005. The three-year measured average PM concentration above background of 2.5 $\mu g/m^3$ for each site within a given air basin was calculated for 2000-2003 and 2003-2005, and the rate of reduction considered. As shown in the following table, our analysis indicated that over the years, PM levels decreased at similar rates across sites within each of air basins examined in California.

Trends in Annual average PM2.5 Above Background, 2000-02 to 2003-05

Basin Name	County	Site	PM2.5 above background ($\mu g/m^3$)		% Change above background since 2000-02 (period2-period1) / period1
			2000-02 (period1)	2003-05 (period2)	
Mountain Counties	Calaveras	San Andreas-Gold Strike Road	6.5	5.3	-19%
	Nevada	Truckee-Fire Station	6.0	4.4	-26%
South Coast	Los Angeles	Lynwood	21.1	16.3	-23%
	Los Angeles	Pasadena-S Wilson Avenue	17.7	14.3	-19%
	Riverside	Riverside-Rubidoux	26.4	20.2	-24%
San Francisco Bay Area	Alameda	Fremont-Chapel Way	9.2	6.5	-29%
	Alameda	Livermore-793 Rincon Avenue	9.8	6.9	-30%
	San Mateo	Redwood City	8.7	6.5	-25%
	Solano	Vallejo-304 Tuolumne Street	10.1	7.5	-25%
	Sonoma	Santa Rosa-5th Street	8.0	5.7	-29%
San Joaquin Valley	Fresno	Fresno-Hamilton and Winery	16.9	14.8	-13%
	Kern	Bakersfield-Golden State Highway	20.4	16.5	-19%
	San Joaquin	Stockton-Hazelton Street	12.8	10.6	-17%
	Stanislaus	Modesto-14th Street	15.2	11.5	-24%
Sacramento Valley	Butte	Chico-Manzanita Avenue	12.1	10.1	-17%
	Placer	Roseville-N Sunrise Blvd	9.9	7.5	-25%

References for Appendix 2

CARB 2004. California Air Resources Board. Staff Report: Review of the California Ambient Air Quality Standard for Ozone, available at:
ftp://ftp.arb.ca.gov/carbis/research/aaqs/ozone-rs/rev-staff/vol4.pdf

Ostro B.C.; Tran, H.; Levy, JI. 2006. The Health Benefits of Reduced Tropospheric Ozone in California. *Journal of the Air & Waste Management Association* **56**:1007-1021.

Appendix 3 (Methodology for Estimating Ambient Concentrations of Particulate Matter from Diesel-Fueled Engine Emissions)

Introduction

This document outlines a method to estimate annual average concentrations of diesel particulate matter (DPM) over large spatial scales. It consists of a simple variation of receptor model, which use measurements of ambient chemical concentrations to infer source contributions, known as the tracer species method.[1] A basic assumption in this method is that the ambient concentration of a tracer species, C, may be used alone to infer the ambient concentration of a pollutant from a specific source, S:

$$S = \alpha C,$$

where α is a scale factor that is independent of location. In the estimation of DPM, we take C to be the ambient concentration of NO_x and S to be the ambient concentration of DPM less than 2.5 µm (DPM2.5). The factor α relates the concentration of PM produced by diesel-fueled engine emissions to the concentration of NO_x produced by all sources. In the following section, we demonstrate that estimates for α based on the emission inventory (EI) and on source apportionment (SA) studies agree within calculated uncertainties. We approximate the distribution of α values over counties by a Gaussian distribution with mean 0.023 and standard deviation 0.006 (for the year 2000). This single value for α and associated dispersion may be used to infer DPM concentrations from measurements of ambient NO_x concentrations in all air basins.

Background

The primary interest of the California Air Resources Board in the estimation of ambient DPM concentrations is for assessment of potential cancer risk. For this purpose, annual average ambient concentrations of DPM are needed. These values are used to calculate lifetime average daily doses[2]; multiplication of the average daily inhalation dose over 70 years with a cancer potency factor gives inhalation cancer risk estimates. In previous estimates[3] of DPM10 concentrations, the Air Resources Board (1998) used a method based on ambient total PM10 concentrations. In this approach, one of two factors, rural or urban, which were determined from chemical mass balance source apportionment studies (CMB) and emission inventory estimates (EI), was used to scale PM10 measurement values to obtain estimates of DPM10 concentrations. Air basins that had more or less diesel to total PM10 emissions than the base case had these DPM10 estimates scaled by another factor (that was determined from the EI): the ratio of air basin to base case value of the relative DPM10 to total PM10 emissions. Application of this method, therefore, depends on several elements, the most important of which are: measurements of ambient PM10 concentrations, previous source apportionment work in specific air basins (base cases), and emission inventory estimates. These components are also the primary weaknesses of the method. Specifically, PM10 contains predominantly crustal material and the fraction associated with diesel PM is very small - at most approximately 0.065; early CMB studies may not

be as accurate as more recent organic marker species-based CMB methods; and early emission inventory estimates may not be as accurate in accounting for all source emissions as more recent models. We believe the proposed use of ambient NO_x concentrations is more direct than the PM10 method to estimate DPM concentrations: close linkage of diesel-engine produced NO_x to total emitted NO_x – about half total NO_x emissions are from NO_x from diesel sources – and relatively good correlation of ambient with recent emission inventory estimates for α.

Methods

In this section, we develop an approximate value for α, the ratio of ambient DPM to NO_x concentrations. First, we compare the ratio of ambient concentrations DPM/ NO_x from several source apportionment (SA) studies with the ratio of annual emissions $(DPM/ NO_x)_e$ from the 2000 emission inventory (EI). Currently, the source apportionment studies are considered the best available methods for determining ambient DPM concentrations (at selected monitoring sites); agreement between the SA and EI estimates of α is used to support the use of a single α value for the whole state of California. Second, based on this favorable comparison, we use the distribution of county EI estimates for the $(DPM/ NO_x)_e$ to determine an average and standard deviation for α.

In the following, we estimate the ratio of DPM to NO_x concentrations for ambient air for two year-long and several short-term source apportionment modeling studies with co-located NO_x measurements. These studies utilize organic chemical speciation for chemical mass balance (CMB) apportionment of PM, which is considered to be essential for the accurate separation of gasoline from diesel-fueled engine emissions. A substantial source of uncertainty in all these studies, however, is in the off-road diesel source contribution. These sources are captured by CMB modeling only to the extent the emissions are similar in chemical composition to those of on-road diesel trucks. In light of the emission inventory estimate that approximately half the diesel contribution to PM and NO_x is from off-road sources, this poorly understood aspect of SA modeling warrants qualifications in all CMB estimates of DPM.

The first considered year-long PM source apportionment work was part of the Children's Health Study (CHS 1995), in which James Schauer carried out organic chemical PM CMB studies for 11 sites in the South Coast Air Basin.[4,5] Hence, 11 annual average values for DPM and NO_x concentrations are available from this work. Two of the sites are centrally located (North Long Beach and Riverside), while the rest are in more or less outlying areas. The second considered SA study was carried out as part of the Central Regional Particulate Air Quality Study (CRPAQS 2000) by Desert Research Institute (DRI) in the San Joaquin Valley.[6] From this work, 6 estimates of annual average DPM and associated NO_x are available. Most of these sites are in urban areas (with the exception of Bethel Island). Although J. Chow of DRI used a different methodology to measure elemental and organic carbon (IMPROVE method) than used by J. Schauer for CHS (NIOSH method), DRI utilized similar specific organic chemical markers for combustion sources. In addition to these long-term measurements, side-by-side CMB modeling was done at two sites for one week each in southern California in

1999 by the two foremost organic marker CMB modelers, E. Fujita and J. Schauer, as part of the Diesel-Gasoline Particulate Split Study (2000).[7,8,9] An unexpected result from this study is that apportionment of PM depends on the specific carbon measurement method utilized (to determine relative organic/elemental carbon). Such differences in apportionment are currently not incorporated into uncertainty estimates. We also note that the Diesel-Gasoline Particulate Split Study raised several important, but still unresolved, questions in the interpretation of CMB modeling results. Specifically, SA estimates may be very sensitive to the choice of source profiles used; e.g. the characteristics of the "average" driving cycle, categories of vehicles, composition of the fleet (e.g. inclusion of high emitter categories such as gasoline "smoker" vehicles) and, information about average high emitter organic species emissions. These aspects bear directly upon SA attribution estimates in a poorly understood manner. Results from several recent short-term apportionment studies that do not utilize CMB modeling are also included below; these studies provide further evidence for a wide range of DPM estimates. Based on a comparison of SA and EI results, we develop an estimate of the DPM/ NO_x ratio from the EI.

Results

Figure 1 shows site-to-site variation of source apportionment estimates of the ratio (annual average DPM10 concentration)/(annual average total NO_x concentration) from the CHS (1995). The sampling sites are described in the CHS Final Report and represent 11 communities in the South Coast Air Basin; these include four urban sites, two sites in a mountainous region, one desert site, three rural coastal sites, and one rural inland site. NO_x measurements and filter samples (organic chemical marker measurements) were taken at the same locations. A straight average over all 11 sites of the ratio DPM10/ NO_x, gives the mean value as 0.024 (0.011), where here and in the following the value in parentheses denotes the standard deviation. An alternative estimate based on regression of DPM10 concentrations against ambient NO_x concentrations (over 11 sites) gives 0.022 (0.009); see Fig. 2. In this, and all following regressions, the intercept is set to zero, which makes the regression less sensitive to scatter and is physically meaningful, as one expects that diesel emissions tend to zero with total NO_x emissions. Removal of an outlying value (for Mira Loma) gives a slope of 0.026 (0.006), which is also shown in Fig. 2.

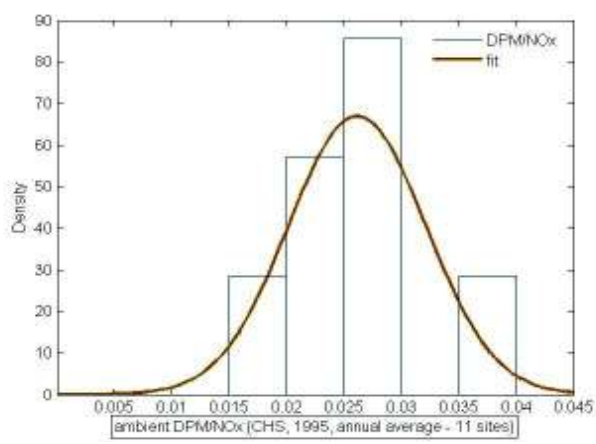


Figure 1

As expected, the dispersion in α is much larger over individual measurements of DPM/ NO_x than it is for the regression coefficient. It is unclear which choice of error is best for use in personal exposure estimates that use population weighting. The site-specific DPM/ NO_x values, Fig. 1, are best estimates for local DPM/ NO_x ratios, though specific meteorology and lack of population weighting may emphasize unrepresentative values. Similarly, DPM/ NO_x ratios obtained from linear regression (with zero intercept) are highly influenced by data with large NO_x and/or DPM values. Because ambient NO_x concentrations may not be related to population density, we believe the statistics for the ratio DPM/ NO_x are better estimates than regression coefficients for DPM exposure-related work. We take the standard deviation of the distribution as the measure of uncertainty in α for SA studies.

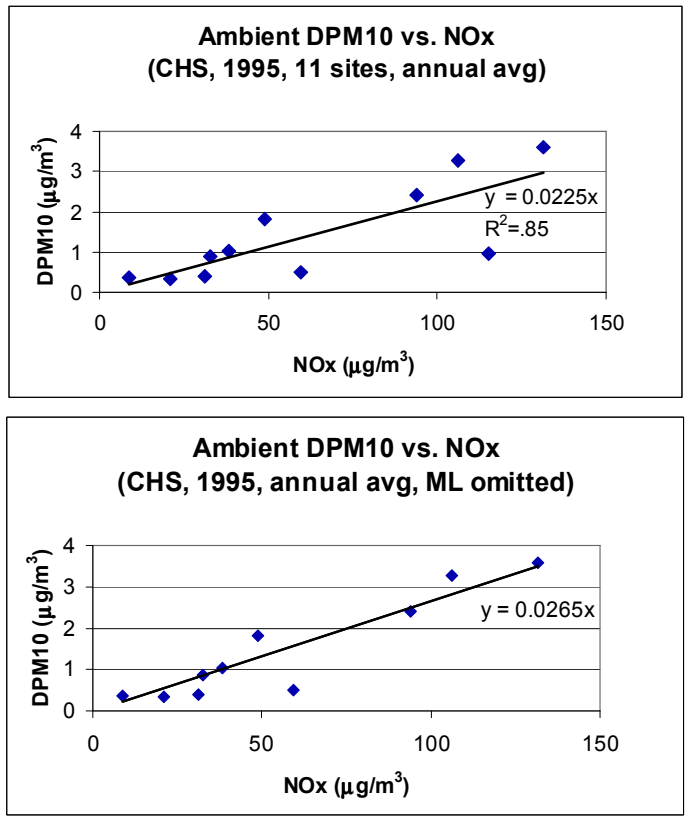


Figure 2

The other year-long SA estimate for α is from CRPAQS (DRI, 2000) for the San Joaquin Valley. A straight average of the ratio of SA DPM to NO_x concentration for 6 sites in SJV gives 0.017 (0.009). Figure 3 shows a regression of SA ambient DPM against NO_x, which gives a slope of 0.015 (0.004). As for the previous SA work, we take the standard deviation (0.009) from the distribution of DPM/ NO_x values as an indicator of the variability in ambient ratios.

We note that the relative variability of DPM/ NO_x in both studies is very large: standard deviation/average ≈ .5 (.011/.024, .009/.017). We believe this large uncertainty in SA estimates best captures local variation of source composition, mixing, chemical reactions and other factors. Hence, this order of uncertainty is expected in any estimate of DPM based on ambient NO_x concentrations.

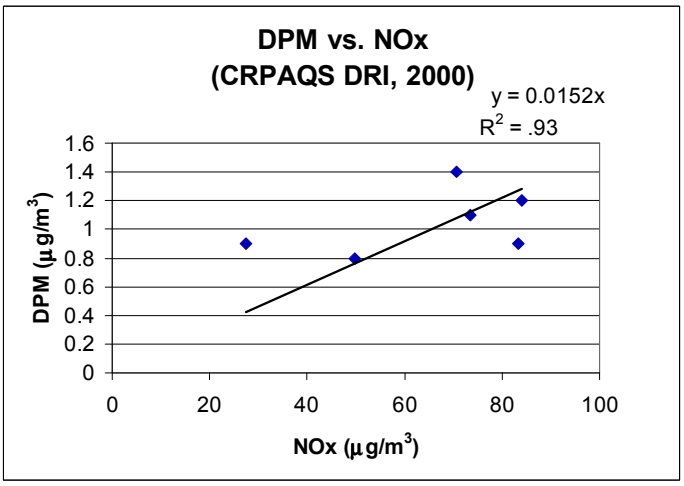


Figure 3

A recent short-term SA modeling

study investigated the sources of uncertainties in the relative contributions of diesel and gasoline vehicle emissions to PM2.5 in the south coast (2001) – the Gasoline/Diesel PM Split Study.[7,8,9] In this work, James Schauer (University of Wisconsin, Madison) and Eric Fujita (Desert Research Institute) collected samples side-by-side for sources (57 light duty gasoline and 34 heavy duty diesel vehicles) and ambient air (two sites Los Angeles, N. Main, and Azusa), and carried out independent chemical and SA data analyses. The SA results show a lack of agreement between diesel PM estimates: apportionment of PM to diesel emission sources by the two groups differ by approximately a factor of two; see Fig. 4. Estimates for DPM2.5/ NO_x are: .010 (.003) Schauer and .023 (.004) Fujita. Because ambient and vehicle emission samples were collected side-by-side, these results indicate that the disparity in DPM estimates are driven by differences in SA methodology, which includes differences in carbon measurement methods (NIOSH vs. IMPROVE), organic marker chemical species, and chemical marker profiles for vehicles. Without *a priori* information about which method is more accurate, we believe both estimates should be weighted equally, giving DPM/ NO_x = .0165 (.009).

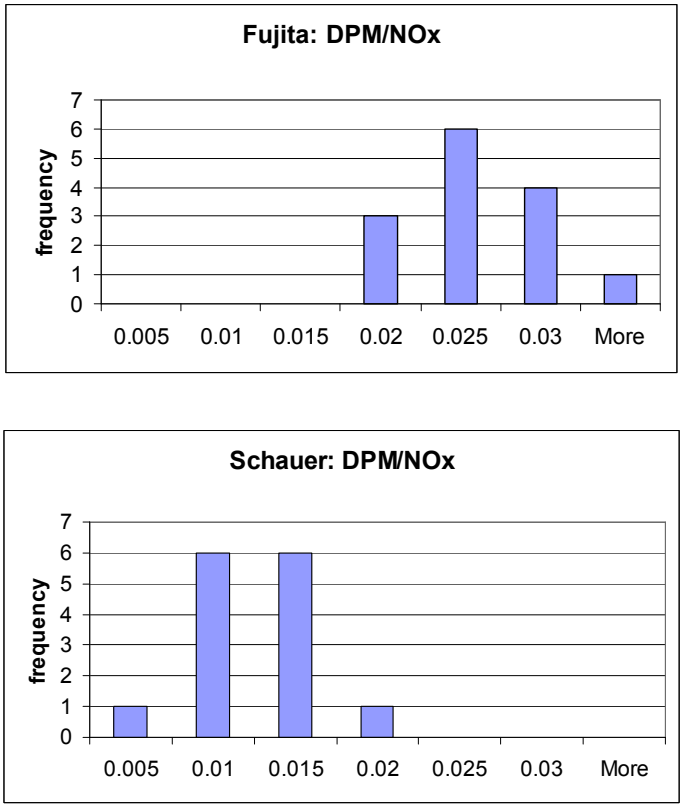


Figure 4

Recent analyses of ambient PM by Livermore National Laboratory (LLNL) in 2007 and ARB's Monitoring and Laboratory Division (MLD) in 2003 gave estimates of DPM concentrations that are similar to J. Schauer's, but not E. Fujita's, results for the Gasoline/Diesel PM Split study: DPM concentrations on the order of 1 μg/m^3 (precise estimates and analyses with colocated NO_x measurements await further work). These values would presumably support the lower DPM/ NO_x ratio of .01 (with a likely relative uncertainty of 50 percent). These studies used methods other than CMB to apportion PM to diesel sources: LLNL utilized fossil carbon measurements (based on Carbon 14) and MLD utilized n-octadecane as a tracer. LLNL show that the average fossil elemental carbon (FEC) at Wilmington is approximately 1.05 μg/m^3 (based on the limited data), and the average FEC at Roseville is approximately 0.65 μg/m^3, which, assuming that all FEC is from diesel emissions and that OC emissions from diesels are small in comparison, may be considered upper bound DPM concentrations. MLD's study yielded estimates of DPM for Wilmington as 1.2 μg/m^3 and Sacramento as 0.8 μg/m^3, and the statewide average as 1.0 μg/m^3. These estimates, however, differ by

over a factor of 2 from the recent MATES III organic marker CMB estimate of more than 3 μg/m^3 in 2004-2005 (in Wilmington). Therefore, while these two independent estimates, yielding approximately 1 μg/m^3 ambient diesel PM (in the South Coast air basin), provide further support for the lower end of DPM/ NO_x ratio, considerable uncertainty remains (CHS, Schauer's Diesel/Gasoline PM Split, and MATES III support higher DPM concentration estimates).

A comparison of the above SA estimates with the emission inventory can not be made directly: emission inventory estimates are for whole counties while SA estimates are specific to monitoring sites and implicitly take into account meteorology, chemistry and deposition. Hence we compare average values for DPM/ NO_x from the previous SA studies with EI estimates of DPM to total NO_x emission ratios. For this purpose, the EI estimates for DPM and total NO_x emission rates for individual counties are utilized.[10] These estimates may be visualized as tons of pollutants emitted each day into a well mixed box covering each county, with removal rates of DPM and NO_x proportionately the same. The assumption of equal removal rates is difficult to verify, given that the rates are caused by deposition, chemical reactions, and flow out of air basins. Further, while the atmospheric lifetimes for DPM and NO_x are typically very different (greater and less than a few days, respectively), which would bias the ratio of DPM/ NO_x toward higher values, the mean residence time of an air parcel in a coastal air basin is often a few hours, which would greatly lessen the difference in removal rates. Given this rough basin lifetime, we assume in the following equal removal rates for NO_x and DPM, and take the overall agreement between SA and EI estimates as support for this assumption.

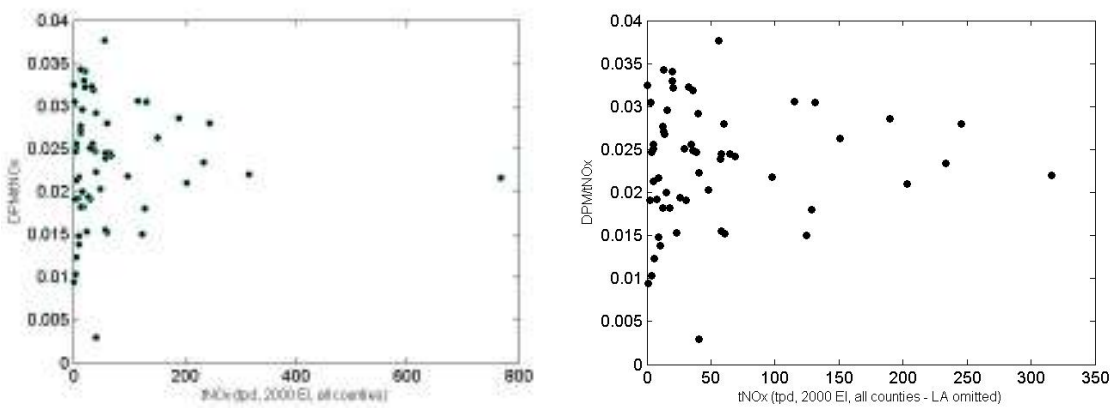


Figure 5

Emissions inventory estimates incorporate spatial and temporal averaging over large scales and therefore may be used to estimate average ambient DPM/ NO_x ratios directly (in this and following expressions, we abbreviate total NO_x by NO_x alone). A plot of (DPM/ NO_x) against NO_x for all counties in California is shown in Fig. 5. Omission of Los Angeles county, which contributes an extremely high value of NO_x (average tons per day), results in the second plot in Fig. 5. These scatter plots show that the ratios DPM/NO_x are clustered about an average and that the dispersion depends on the

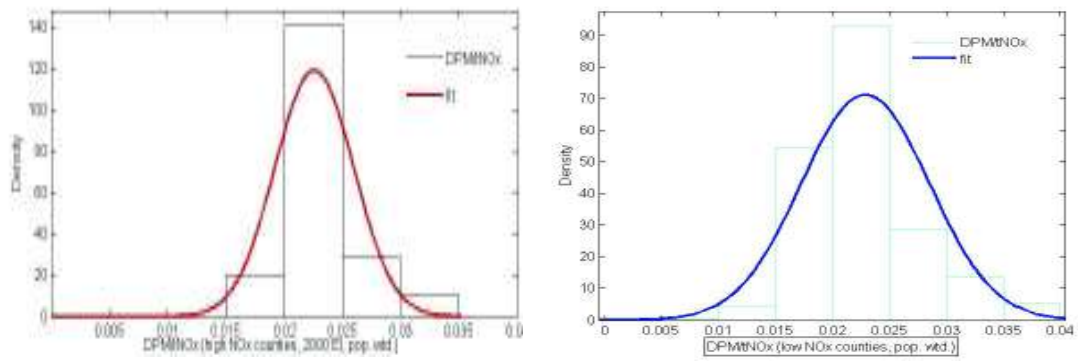


Figure 6

average annual NO_x emission rate. The second plot in Fig. 5 shows that a separation of high-NO_x from lower-NO_x emission counties occurs with a division around an annual average of 80 tons per day. (High- NO_x counties are listed in the section Results.) High-NO_x counties are highly urban and have similar composition of diesel to non-diesel emission sources. To better capture exposure-related estimates of DPM/ NO_x, each county value is weighted by its population; weighted histograms are approximated by normal distributions. Figure 6 shows the high- and low-NO_x emission distributions for α. The mean and standard deviation for α are: 0.023 (0.003) for the high-NO_x county estimate and 0.023 (0.006) for the low-NO_x county estimate. In summary, the distribution for the factor α is described by a single mean value, independent of high and low-NO_x counties, and a dispersion that depends on whether the county is highly urban or not. We take the greater deviation of DPM/ NO_x ratio for low- NO_x counties as measure of the variability that is locally encountered within air basins.

The above estimates of the ratio DPM/NO_x from the EI-population weighted and SA studies compare well, given the relatively large uncertainty: EI: county average 0.023 (0.006); and SA: SC 0.026 (0.006), SJV 0.017 (0.009), and Gasoline/Diesel PM (GDPM) Split SCAB 0.017 (.009); see Fig. 7. This agreement between EI and SA estimates for α motivates the use of a single scaling factor for the whole state of California to estimate annual average concentrations of DPM from annual average measurements of NO_x. We take the EI

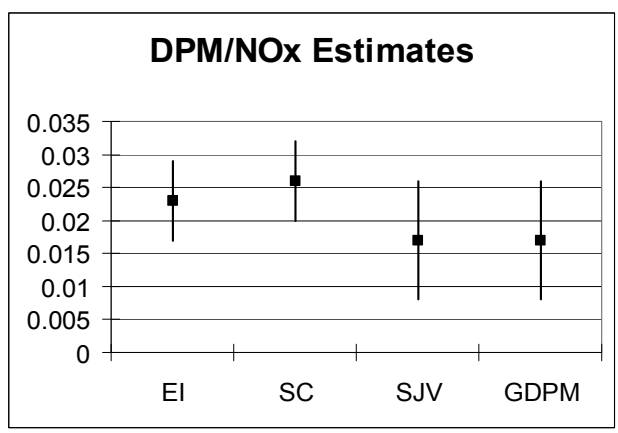


Figure 7

values for the average and standard deviations low-NO_x emission counties as best estimates for DPM/ NO_x; these counties capture some of the variation in emission sources that is encountered locally. The value α = 0.023 (0.006) is a population weighted estimate of DPM/NO_x for all locations in California; the standard deviation indicates the uncertainty in this choice of α for a given county.

Conclusions

Based on the agreement between SA and EI estimates of the scaling factor α, the ratio DPM/total NO_x, we propose the use of a single value of α for estimating the population-weighted annual average ambient DPM concentration for California. These DPM estimates depend upon the network of ambient NO_x measurements from the ARB monitoring sites. In the following, we outline a method to calculate such averages. First, the annual average DPM concentration at each monitoring site is estimated as the product of annual average NO_x concentration value and α. The uncertainty associated with this DPM estimate is the product of the annual average NO_x measurement value and the low-NO_x county standard deviation, .006. [Although not utilized, the following twelve counties are considered high-NO_x (annual average NO_x more than .80 tons per day): Los Angeles, San Bernardino, Kern, San Diego, Orange, Riverside, Alameda, Fresno, Santa Clara, Contra Costa, San Joaquin, and Sacramento; the remaining 46 counties are considered low-NO_x counties.] From this set of spatially discrete DPM concentration estimates a smooth DPM concentration surface may be constructed using kriging or other methods. In remote areas without monitoring sites, the smoothing method may be modified to incorporate a minimum concentration, which reflects a nonzero background value (or such areas may be removed, if the population is sufficiently small). Second, census data for California is used to approximate a population density surface (population fraction per unit area) and the product of the population density and DPM concentration surfaces (pointwise) is taken. This product may be integrated over any region and divided by the fraction of California population within that region to give a population-weighted average DPM concentration; in particular, integration of the product may be performed over the state to give an average population-weighted ambient DPM concentration. Once ambient diesel PM concentrations have been estimated for a baseline year (2000), linear rollback techniques may be used to project concentrations for future years.

DPM concentration estimates ($\mu g/m^3$)			
Air Basin	Population	Previous	Proposed
Great Basin Valleys	32006	0.1	0.18
Lake County	58309	0.2	0.54
Lake Tahoe	46200	0.4	0.24
Mojave Desert	816742	0.1	1.46
Mountain Counties	408039	0.1	0.43
North Central Coast	710598	0.8	0.59
North Coast	310061	0.8	0.33
Northeast Plateau	87578	0.7	0.18
Sacramento Valley	2334277	1.3	1.02
Salton Sea	465886	1.5	1.29
San Diego County	2813833	1.4	1.49
San Francisco Bay	6605921	1.6	1.62
San Joaquin Valley	3189385	1.3	1.36
South Central Coast	1399218	1.1	0.93
South Coast	14592351	2.4	2.90
Statewide (pop. wtd.)	33870404	1.8	2.00

Table 1

A comparison of DPM concentration estimates for the year 2000 from the proposed NO_x-scaling method (proposed) with the projections from the previous PM10-scaling method[3] is given in Table 1. The overall agreement between DPM concentration estimates is good, and for the six highest population air basins is very good. More specifically, the six highest population air basins contain over 90 percent of the population of California and contribute greater than 96 percent of the population weighted DPM concentration; in each of these air basins, the difference between the proposed and the previous DPM concentrations is less than approximately 20 percent (of the previous estimate). It should be noted that the previous estimates use a baseline year 1990 and are projected forward by a decade based on linear rollback, and so do not constitute the best approximation for year 2000. Greater variation of agreement between proposed and previous methods is found for lower population air basins. Many factors contribute to this variability, several of which are: the larger dispersion in the DPM to NO_x ratio (.006), uncertainty in application of PM10 scaling method to regions less similar to the SJV, and greater influence of localized emission sources. Altogether, the proposed, population-weighted DPM concentration for California is increased by 11 percent over the previous estimate. This high level of agreement between the population-weighted DPM estimates gives confidence that the proposed method is consistent with the previous technique and represents a viable approach to estimate DPM exposure.

In summary, the proposed method to estimate ambient DPM concentrations has distinct advantages over the previous PM10 method as well as several important limitations. The primary strengths of the method include the strong relation of DPM to (total) NO_x; simple application; estimates of uncertainty intervals; and ability to capture sub-county variations in DPM concentrations. In addition to these strengths, the approach is tied directly to the ARB emission inventory, and links bottom-up EI estimates with top-down SA estimates. Several limitations and caveats also bear on applications of the method. The limitations include all assumptions sufficient for application of EI estimates to ambient air, such as well-mixed air parcels (county scale), proportional removal rates for NO_x and DPM, proportionally uniform emission rates for all NO_x and DPM sources, etc. Verification of these assumptions is in general not possible; instead, agreement between EI and SA estimates is taken as best available evidence. The uncertainty intervals produced by the estimation method are based on variations between low- NO_x counties and reflect differences in relative emission sources (primarily diesel vs. non-diesel). As such, the uncertainty describes the confidence in α to accurately describe local NO_x emission sources. Further work is needed in strengthening the understanding of the contribution of various emission sources to ambient concentrations of both gases and particles. In this respect, source apportionment work that utilizes organic marker species is the best available approach; ideally, highly time-resolved studies would allow better characterization and support for single species scaling estimates, such as the NO_x-scaling method. Finally, off-road diesel sources, which are a large source of uncertainty in current CMB modeling, need to be explicitly included in future source apportionment studies (i.e. chemically characterize emissions as a function of operating mode and construct a source profile for CMB modeling work).

References for Appendix 3

1 Henry R., Lewis C., Hopke P., and Williamson H., *Review of Receptor Model Fundamentals* (1984) Atmospheric Environment 18, 1507-1515.

2 *Air Toxics Hot Spots Program Risk Assessment Guidelines, Appendix D Risk Assessment Procedures to Evaluate Particulate Emissions from Diesel-Fueled Engines, 2003,* Office of Environmental Health Hazard Assessment.

3 *Estimate of Ambient PM10 Concentrations from Directly Emitted Emissions from Diesel Engines, Appendix III, Part A, Exposure Assessment from Proposed Identification of Diesel Exhaust as a Toxic Air Contaminant* (1998). Report to the California Air Resources Board.

4 Manchester J., Schauer J., and Cass G. *Determination of the Elemental Carbon, Organic Compounds and Source Contributions to Atmospheric Particles During the Southern California Children's Health Study: Part B; The Distribution of Particle-Phase Organic Compounds in the Atmosphere and Source Contributions to Atmospheric Particulate Matter Concentrations During the Southern California Children's Health Study, 1995*. Final Report for California Air Resources Board.

5 Manchester-Neesvig J. B., Schauer J. J., and Cass G. R., *The Distribution of Particle-Phase Organic Compounds in the Atmosphere and their Use for Source Apportionment during the Southern California Children's Health Study* (2003) Journal of the Air and Waste Management Association. 53, 1065-1079.

6 Chow J., Chen L.W., Lowenthal D., Doraiswamy P., Park K., Kohl S., Trimble D., and Watson J. *California Regional PM10/PM2.5 Air Quality Study (CRPAQS): Initial Data Analysis of Field Program Measurements, Final Report, 2005* for the California Air Resources Board.

7 The DOE Gasoline/Diesel PM Split Study, Presentations by D. Lawson, E. Fujita and J. Schauer, California Air Resources Board Seminars webpage; and DOE/NREL Gasoline/Diesel PM Split Study webpage.

8 Fujita E., Campbell D., Arnott W.P., Chow J., and Zelinska B., *Evaluations of the Chemical Mass Balance Method for Determining Contributions of Gasoline and Diesel Exhaust to Ambient Carbonaceous Aerosols* (2007) Journal of the Air and Waste Management Association. 57, 721-740.

9 Lough G. and Schauer J., *Sensitivity of Source Apportionment of Urban Particulate Matter to Uncertainty in Motor Vehicle Emissions Profiles* (2007) Journal of the Air and Waste Management Association. 57, 1200-1213.

10 California Air Resources Board, 2007 Emission Inventory, Planning and Technical Support Division.

Appendix 4 (Peer Review Process and Results)

Cal/EPA has a new Interagency Agreement with the University of California (UC) for External Scientific Peer Review. The Agreement incorporates guidelines for Cal/EPA organizations requesting external review. Reviewer candidates are independently identified by the University of California at Berkeley, Institute of the Environment, in collaboration with UC colleagues.

The request for reviewers to the Cal/EPA Project Director for the proposed methodology resulted in six reviewers being identified and approved for this assignment. Collectively their expertise is based on research in the following areas: chronic obstructive pulmonary disease related to air pollution; statistical analysis of epidemiological data; particle formation and measurements in air; air quality risk management; air pollution and daily mortality associations; and epidemiology. These reviewers received a draft report dated August 23, 2007 and evaluated whether CARB staff correctly interpreted the results published in the literature, including U.S. EPA's expert elicitation, and whether staff correctly developed methods for estimating premature deaths associated with public exposures to ambient PM. The peer reviewers provided staff with written comments on a draft of this report. Staff then addressed and incorporated the results of the peer review into this report.

In addition, the peer reviewers considered the two scenarios and concluded that mortality C-R functions can be applied to small areas and populations, as long as uncertainties and limitations are explicitly stated, including:

- The composition of PM must be limited to sources known to be toxic, such as diesel PM.
- For small populations, the risk can be described as the risk reduced by a certain percentage.
- Demographics of the affected population should reflect the general demographics of the population considered in the original epidemiological studies.
- The concentration of PM should not vary significantly for the population affected.

The peer reviewers also suggested that staff consider that the concentration of PM may vary by an order of magnitude over a distance of 0.5 km, and fine scale modeling may be needed. They also cautioned that the demographics of the small population may not reflect the county population, and adjusting the incidence rate for age, gender and socioeconomic status differences may be needed.

Based on their expertise, two of the peer reviewers were also asked to comment on the proposed methodology for estimating diesel PM concentrations. Their comments are also included in this appendix.

The peer reviewers and their affiliations are:

Jeffrey Brook, Ph.D.
Environment Canada
Adjunct Professor
Public Health Sciences/Chemical Engineering
University of Toronto

Mark D. Eisner, M.D., M.P.H.
Associate Professor
Pulmonary and Critical Care Division
UC San Francisco

Richard C. Flagan, Ph.D.
Professor
Chemical Engineering/Environmental Science and Engineering
California Institute of Technology

Alan Hubbard, Ph.D.
Assistant Professor
Biostatistics
UC Berkeley

Joel Kaufman, M.D., M.P.H.
Professor
Environmental and Occupational Health Sciences
University of Washington

Joel Schwartz, Ph.D.
Professor
Environmental Health/Epidemiology
Harvard University

A. Comments on General Methodology Described in the Draft Report

In this section, a summary of comments from the peer reviewers is presented, followed by individual comments from the six experts.

Summary of Peer Reviewer Comments
On General Methodology Described in the August 2007 Draft Report

Issue	J Brook	M Eisner	R Flagan	A Hubbard	J Kaufman	J Schwartz
Credible Range	10% ok. Upper and lower bounds could be better.	Good.	10% ok. No comment on range.	10% is good.	Good. Should discuss Miller 2007 and newer publications.	Did not fully discuss opinion.
Sensitivity Analysis	Results presented show wider ranges than adopted as credible range. Recommend pooling all 12 expert or 10 expert distributions, but recognize the lower limit of 0 would be problematic.	Delete Jerrett 2005 in one sensitivity run. Pool results of all studies in another run.	No comment.	Consider using sensitivity results to develop upper and lower bounds of credible range.	Do not include both Pope and Jerrett in one run.	Can pool Pope with Jerrett. Point out bias in Adventist study. Add Laden's results on PM change between periods and give Laden more weight.
Cut-off Level	7 ug/m3 is good.	7 ug/m3 is not well-justified. Consider 2.5 ug/m3 as an alternative.	Need to justify dropping 0 ug/m3. Should consider no threshold.	No comment.	No comment.	7 ug/m3 is not defensible. Should use 2.5 ug/m3.
Roll-back	Reasonable. Clarify the use of background 2.5 ug/m3.	No comment.	Revise the formulae and explanations.	Reasonable	No comment.	Revise description for rollback method; as written, it is unrealistic.
Overall	Good.	Good.	Need clarity in several places.	Good.	Good.	Generally good.

A.1 Jeffery Brook

Scientific Review of the Air Resources Board (ARB) Draft Report on "Methodology for Estimating the Premature Deaths Associated with Long-term Exposures to Fine Particulate Matter in California"

Reviewer: Dr. Jeffrey R. Brook, Senior Research Scientist, Environment Canada; Adjunct Professor, University of Toronto.

Comments

The comments below focus on the three key components identified in Attachment II. In reviewing the material provided I have considered whether the methodology described represents sound scientific knowledge, judgment, methods and practices. Although knowledge on $PM_{2.5}$ health effects and $PM_{2.5}$ exposure has advanced dramatically in the past 10+ years, understanding of the issue is far from complete. Much remains to be learned about the relative toxicity of different particles based upon their physical and chemical features and how they vary by source and as a result of atmospheric processes. The role of gaseous pollutants in the mix that people breathe and their interactions with and interactive effects with particles also requires clarification. The possibility that the net effect a given particle type can have on health also varies by endpoint (e.g., cardiovascular vs. respiratory mortality) and according to a person's susceptibility is also very real and not well understood. Furthermore, any information we have on these issues has yet to provide a means for more refined concentration-response functions (CRF). Consequently, a significant amount of assumptions must necessarily underlie any method for estimating avoided mortalities associated with decreasing $PM_{2.5}$ concentrations. Above all, this requires scientific judgment, with frank discussion of the assumptions made and the limitations of the method. Overall, the ARB draft report meets these criteria, although below are some comments that may help improve the document and spark some further thinking.

The development of a credible range based on expert opinions from a panel of experts selected by U.S. EPA

ARB's use of the U.S. EPA expert panel process implies two key assumptions:

- That the EPA process was appropriate and rigorous and represents the best approach to developing a CRF given the existing uncertainties, which are exemplified by the range of coefficients found in the different studies considered.
- That all the studies considered by the U.S. EPA (EPA) are relevant for the population and exposure conditions present in California.

In terms of the first assumption, the ARB is fully justified in building on the U.S. EPA's effort for two reasons. Firstly, the EPA's effort was itself thoroughly reviewed and, although there were some concerns expressed by its reviewers, it was deemed to be necessary and of high quality. For an assessment of the CRF relating premature mortality to long term $PM_{2.5}$ exposure, it is unlikely that this effort and its outcome could

be improved upon, given current information.

The second assumption is more difficult to judge due to the limited number of cohort studies of the premature risk posed by long term $PM_{2.5}$ exposure (as represented by an annual or multi-year average ambient concentration). The ARB staff adequately discussed this issue in the draft report. Given the fact that some of the studies were from populations in California, entirely or in part, and the lack of any evidence indicating that the study results are not applicable to California, I find that this assumption is justified.

Therefore, the information used by the ARB staff to develop the low, central and high CRF estimates is appropriate. Among these values, the central estimate of 10% is well explained. Using the median of the medians among all the experts involved in the EPA process is scientifically acceptable. It reflects current knowledge and I do not think that there are any other reasonable approaches that could have been followed. Furthermore, the sensitivity analysis supports this value and so it is well-justified

The values selected for the low and high points in the range are more difficult to assess and the ARB staff pointed out the challenge of determining these points. The question is whether or not the values identified have led to a credible range. From my perspective this is equally difficult to assess since no criteria were provided for what constitutes credible. I will assume here that credible means that there is some science-based evidence to support the range and that the high and low values are reasonable in terms of leaning towards being somewhat conservative and hence not likely to be controversial. Based upon this definition it is my view that the range of 4% to 16% is credible.

There are some important issues that should be addressed in the final version of the report. Firstly, it (the final report) should provide ARB's view of what the high and low values of the CRF actually signify. On page 5 of the report it is stated that they are an uncertainty interval, but is that truly what they are? Perhaps they represent uncertainty in a more subjective manner, but not in the purely objective, quantitative sense that some readers may expect from uncertainty values. Secondly, and related to the first point, the final report should provide a discussion of how ARB would use (i.e., communicate) results calculated from the upper and lower limits. Given how they were determined, it does not seem, as indicated above and below, that truly they express the degree of uncertainty about the central estimate. These comments are somewhat outside pure scientific review, however, selecting the range involves both objectivity and subjectivity and thus, it is important to clarify what the purpose or meaning of those values is expected to be. Ultimately, that is the only way to guide their quantification and application.

In the final report the way in which the upper and lower values in the range were determined needs to be explained in more detail to assist readers in assessing their scientific credibility. The general concept of bounding the range based upon the larger value from the "Six Cities follow up" and the lower value from the ACS is clearly

described. The reason for doing this is that the ARB staff speculated that developing the upper and lower bound from the full spectrum of expert opinions may be highly influenced by their "high" and "low" opinions. This may be possible, however, the full outcome of the expert solicitation should not be taken lightly. In their independent and collective deliberations they were equally aware of which studies were the key ones (i.e., ACS and Six Cities) and which ones could inform the possible range or uncertainty. In the draft report it is stated in the middle of page 27 that "Staff chose to rely on empirical evidence to bound the central estimate." I assume that what is meant by "empirical evidence" is that the result of a single study is considered to be empirical because it was purely a quantitative, statistical analysis, as opposed to expert opinion. The final report should clarify this and indicate exactly how 4% and 16% were obtained.

The upper bound of 16% appears to be in Table 1 (directly from Laden et al.). This value is further supported as being a plausible based upon the recent ACS L.A. sub-study (Jerrett et al., 2005). However, both of these studies (i.e., Laden et al. and Jerrett et al.) had upper confidence limits of 26-30% and so choosing the risk coefficients obviously is not recognizing the full range of uncertainty found in that research. Thus, a key point to realize is that ARB's recommended upper bound is smaller than the upper confidence limits of some of the studies and of some of the expert panel member's opinions. Thus, ARB has leaned towards being conservative on this issue. This is a prudent choice and any impact or benefit calculations using the upper bound should be less likely to be controversial. The final report should consider pointing this out.

The lower bound is potentially more controversial. It is also not clear where 4% came from based upon information in the figures or on Table 1. Thus, as indicated above, the final report needs to expand the middle paragraph in page 27 with more specifics. More about the lower bound will be discussed in the next section on sensitivity analysis.

Sensitivity Analysis

This analysis is important due to the lack of a single best approach to determine upper and lower bounds (i.e., the credible range) and the central or mean CRF. It helps support the values proposed by the ARB. Given the available information, the method developed by the ARB staff is scientifically acceptable in that multiple approaches were considered and evaluated against the recommended values. However, it is noted that ARB's range is narrower than any of these approaches. For the upper end, this implies that ARB is being conservative, but this is not the case in the choice of a larger lower end.

One difficulty from the results of the sensitivity analysis and from the range recommended by ARB is that any of the seven approaches included in the sensitivity analysis could probably be rationalized as being a credible approach. Overall, the most objective ones are probably #6 and #7 as they essentially remove ARB staff from the equation. If credible scientists rigorously polled highly reputable experts and other experts carefully reviewed the process (i.e., EPA's expert elicitation), then why not let that process speak for itself (i.e., used #6 or #7 to get the range)?

Although it is hard to follow how the draft report's description of what the random effects approach is supposed to account for (i.e., that the different values may have come from different distributions due to there being different CRFs potentially because of varying $PM_{2.5}$ composition) justifies its use for pooling expert opinions, the bottom line is that it is probably a more conservative approach than just taking a variance-weighted average. However, the challenge is that Table 3 shows that a lower limit of zero was obtained. There is a big difference between zero and 4% (the lower bound selected by ARB). Thus, the final report needs to provide a reason for the lower limit being positive and why that is more credible. I suggest that there is more than enough *in vivo* and *in vitro* toxicological data and human clinical data (i.e., biological plausibility) to support the notion that $PM_{2.5}$ does have an effect. Thus, it is highly likely that the lower bound is not zero and the evidence for this is much greater today than 10 years ago. Furthermore, given the tendency for the more recent cohort analyses and intervention studies to yield larger effects than the earlier work probably supports the larger lower range (i.e., 4%) compared to the other non-zero lower bounds derived from the sensitivity analysis.

The overall picture is that I do feel that sensitivity analysis provides some added and valuable scientific rigor to ARB's work, it was reasonably well done and it helps support what I agree to be a credible range of 4-16%.

Estimation of premature death associated with exposures to $PM_{2.5}$

The approach ARB proposes to use is discussed on pages 30-34. My opinion is that what is proposed is based upon sound scientific knowledge, judgment, methods and practices. Where possible, units should be stated for the variables in the equations (Y_0 and β). The available $PM_{2.5}$ data are used appropriately to estimate the population exposure. Although the interpolation method used to assign monitoring site $PM_{2.5}$ concentrations to census blocks is relatively simple and does not consider terrain features or prevailing meteorological features that might distribute the particles differently across the state, it would require considerably more work to gain any improvements. Newer approaches such as land-use regression or data fusion are currently beyond the scope of the current draft report. ARB should check the maps in Appendix 1. The interpolation and contouring results for the latter two years and for the far SE portion of the state look different that I would expect given the concentrations around the nearest monitoring sites. Clearly, this would have little impact on any results.

Three cut-off levels, below which there are no benefits (avoided mortalities) to further reductions in annual average $PM_{2.5}$, were discussed in the report. Given the lack or information regarding the true value, if one exists given the ranges of susceptibility in the population and the possibility that it would be different for different endpoints or causes of mortality, the proposed value of 7 $\mu g/m^3$ represents sound scientific judgment. I agree that 2.5 $\mu g/m^3$ is too low and there are not sufficient data to adequately evaluate if annual average $PM_{2.5}$ levels between 2.5 and 7.0 $\mu g/m^3$ are associated with changes in mortality rate or whether or not β is different in this range. However, using a value as low as 7 $\mu g/m^3$ as opposed to 12 $\mu g/m^3$ is well-justified based upon the ACS range and Pope et al.'s findings. Furthermore, time series studies indicate that there are acute mortalities occurring in communities with annual averages less than 12 $\mu g/m^3$. Thus,

this value is clearly too high.

To better understand the impact of these different cut-off values the ARB may want to consider future sensitivity studies where the number of avoided mortalities due to a proposed policy or a roll-back to attainment is computed using each of the values and then are compared. In the context of the types of changes in emissions to be expected via new policies on "goods movement", it seems unlikely that the use of 2.5 or 7 $\mu g/m^3$ for the cut-off would make much difference. However, using different values between 7 and 12 $\mu g/m^3$ could affect such results.

In the second part of this section of the draft report, where ARB describes how to determine ΔPM given the max concentration in a basin and the cut-off value, there is one key assumption. That is that any roll-back strategy (i.e., the emissions reductions to attain the standard) to get the BasinMax into attainment will proportionately affect all other $PM_{2.5}$ monitoring sites and hence the population exposures within the basin. This is a reasonable assumption for crude roll-back analyses and, in general, data in the Appendix support it. However, in the context of the types of changes in emissions to be expected via new policies on "goods movement" this assumption would not likely hold. Clearly, ARB must be aware of this fact and would be constructing much more detailed base case and future case exposure maps under different policy scenarios. Finally, in this part ARB has set BG=2.5 $\mu g/m^3$. It is not clear to me if this is where the new cut-off value would be used. If this is the case, then I presume that 2.5 $\mu g/m^3$ is a "typo". If this is not the case then where and how does the cut-off value enter into the estimation of avoided health impacts?

Final Comments

The draft report and the methodology described are scientifically sound given current information on $PM_{2.5}$ health effects. The range for the CRF is credible and reasonably conservative and, as pointed out in the draft report, the true benefits that can be ascribed to reducing $PM_{2.5}$ are likely to be larger still because of endpoints that currently cannot be quantified. There are parts of the draft report that would benefit from some clarification and additional discussion, as noted above.

A.2 Mark D. Eisner

Critique of "Methodology for estimating the premature deaths associated with long-term exposures to fine airborne particulate matter in California." ARB, California EPA.

Mark D. Eisner, MD, MPH
UCSF

1. DEVELOPMENT OF A CREDIBLE RANGE BASED ON EXPERT OPINION

The elicitation process used by U.S. EPA and adapted by this report is robust and appropriate.

The issue of geographic appropriateness regarding the health effects estimates for $PM_{2.5}$ was discussed on page 24. One issue to consider is potential interactions between SOx, ozone, and $PM_{2.5}$. Because ozone and SOx levels vary geographically, would the health effects of PM differ in California vs. other areas with different ozone and SOx levels?

2. SENSITIVITY ANALYSIS

The use of the ACS and Six Cities studies to develop the upper and lower uncertainty limits does not take into account the variability around the risk estimates from each study (i.e, the 95% confidence intervals). The authors should consider an additional sensitivity analysis in which the lower 95% CI bound of the ACS and the upper 95% CI bound of the Six Cities studies are used. This would better reflect the variability implicit in those estimates.

On page 27 it is stated that it is technically incorrect to pool non-independent results from the same underlying cohort study (i.e., Pope 2002 and Jerrett 2005). It is therefore difficult to understand why it was done. The effect is to give greater weight to the ACS study. Consideration should be given to deleting the Jerrett analysis from the sensitivity analysis.

A suggestion for an additional sensitivity analysis would be to pool the results of all studies that measure $PM_{2.5}$ and all cause mortality, even those that have issues of genearlizability to the overall California population (e.g,. ASHMOG). The inclusion of non-generalizable studies would appear to be a less serious issue that the inclusion of more than one analysis of the same study (i.e., non-independence).

3. ESTIMATION OF PREMATURE DEATH

Estimation of PM concentration. It is stated on p.30 that there may be no monitoring information within 50 km. More information should be provided about what proportion of census blocks for which this is true. A sensitivity analysis excluding these centers should be considered to evaluate the impact of these centers on the effect estimates for

$PM_{2.5}$ and mortality.

Estimation of the mortality impact (p.30). The equation indicates a Beta coefficient. One presumes that this is for a 1 ug/m^3 $PM_{2.5}$ increment, but this should be clarified. In addition, there is a discrepancy between the baseline death rates, which includes all deaths over the entire population of all ages, and the "pop" variable which includes the population aged 30 years or greater. Can the baseline death rate and population variables be based on the same age ranges?

The issue of a $PM_{2.5}$ cut-off value. The analysis uses a cut-off $PM_{2.5}$ value of 7ug/m^3. Yet it is stated that 11/12 experts agreed that health effects may be observed at all levels of $PM_{2.5}$. The proposed analysis defines all exposure less than 7ug/m3 as zero exposure. This does not seem appropriate given the lack of evidence for a threshold effect. At a minimum, an alternate analysis that allows for linear extrapolation down to the background level of 2.5ug/m^3 should be performed.

On page 37 the statement is made that "Although the literature mostly favors a no-threshold model, without empirical evidence for PM effect between 2.5 and 7ug/m^3 we recommend that no premature deaths be associated with PM exposures in this range. As discussed above, this seems illogical. Although the functional form of the relationship between $PM_{2.5}$ and mortality in this range is not known, assumption of a linear relationship would appear to be more sound than to assume no health effects at all.

There are no results presented for the roll-back analysis. The methodology is presented, but the results are not.

A.3 Richard Flagan

Review of Proposed Methodology to Estimate Premature Deaths Associated with Long-Term Exposures to Fine Airborne Particulate Matter in California. (R. Flagen)

The methodologies described in this report are based upon results of a series of epidemiological cohort studies that provide an empirical basis for estimating premature deaths associated with exposure to fine particulate matter. At the same time, the challenges faced by the researchers who performed those studies raise fundamental questions about strategies for monitoring air quality, and that limit the resolution of the statistical analyses. The studies that were ascribed the highest reliability by the experts consulted in the EPA study employed $PM_{2:5}$ measurements of atmospheric, fine particle mass concentrations. Decades of such measurements at community monitoring stations in a number of cities have enabled the development of the methodology outlined in this report. Recent literature raises serious questions that suggest that $PM_{2:5}$may just be the tip of the iceberg - that associations with smaller particles should be explored, but the data for such proactive studies neither exist nor are likely to become available in the near future.

Traditional aerosol exposure monitoring reports only mass concentrations in a few broad size ranges: PM_{10} - particles smaller than 10 μm in diameter (Dp<10 μm), and $PM_{2:5}$ - fine particles for which Dp<2.5 μm. Exposures to fine particles are associated with a range of health consequences (Pope and Dockery, 2006) from increased asthmatic symptoms (McConnell et al., 1999) to decreased lung growth (Gauderman et al., 2000, Gauderman et al., 2002) to mortality (Pope et al., 2002, Jerrett et al., 2005). Mass based PM_{10} and $PM_{2.5}$ measurements are, for several reasons, blunt instruments for the assessment of exposures to potentially harmful particulate matter. Within any size fraction, the mass concentration is biased to the largest particles in the included size range. Numerous studies provide evidence that particle mass is not the best measure for potential health effects of fine particles, and that the smallest particles in the fine particle size fraction may have the most profound health effects (Oberdorster, 2000; Donaldson, et al. 2002). These effects cannot be found in epidemiological studies because the vast majority of air quality measurements are limited to those parameters that are covered in present regulations. This is a fundamental failing of the present air quality monitoring system. Until air quality monitoring goes beyond the presently regulated quantities, it will remain impossible to develop health effect associations with suspected, but unregulated (and hence unmeasured) atmospheric contaminants.

A more effective partnership between epidemiologists and health researchers, atmospheric scientists, and regulatory agencies will be required if emerging health problems are to be identified without decades of delay as fine particulate matter health impacts have required. This will require investment in the measurement infrastructure in addition to acquisition of health-related atmospheric exposure data. Instruments need to be developed that can provide data on contaminants of interest that meet the stringent needs of epidemiological studies, especially the ability to provide robust data at a cost that is compatible with extended duration, large scale studies. Lacking such foresight,

future attempts to assess health impacts will, like the present studies, be forced to rely on studies that do not fully constrain the exposure assessments.

The present methodology document does not address the questions raised above, but rather works within the constraints of the existing air quality and epidemiological data. In the discussion that follows, I have focused my comments on three basic questions that arise from the proposed methodology.

> *Question 1: Does the methodology in the present report provide a rigorous basis for the new relationship for estimating premature deaths associated with long-term exposures to fine particulate matter in California?*

The methodology is based upon a careful review of the relevant literature; with emphasis upon the studies that are most widely accepted for provide the best quantitative estimates for the prediction of premature death rates. The data employed in those studies is limited, as outlined above, and some of the studies did not even have the full $PM_{2:5}$ data. In spite of the atmospheric data challenges, the studies produce a remarkably consistent picture of the effects of fine particle exposures. The methodology development study has also consulted EPA expert evaluations of the previous studies, which involved interviews to elicit assessments from 12 world-renouned experts on health effects of air pollutants. The CARB analysis of those studies considered subtle factors that might have influenced the EPA recommendations, and provide a clear basis for the recommendation that the relative risk of exposure to PM2.5 be a 10% increase in premature death rate per 10 $\mu g/m^3$ increase of $PM_{2:5}$ exposures.

> *Question 2: Does the methodology provide a reasonable basis for the assessment of the threshold for the effect of $PM_{2:5}$ exposure on the premature death rate?*

Here, I have difficulty in understanding the rationale presented for the premature death rate. The report notes that the suggested threshold of 7 $\mu g/m^3$ corresponds to the lowest levels observed in the Pope et al. (2002) study. Eleven of the twelve experts consulted by the EPA discounted the idea that a threshold exists in the influence of $PM_{2:5}$ on the premature death rate. The experts who favored epidemiological studies for determination of threshold effects conceeded that definitive studies needed to ascribe a threshold would be difficult or impossible.

In their considerations for the present methodology report, the CARB considered three alternatives for a threshold value, 2, 12, and 2.5 $\mu g/m^3$. No justification is provided for excluding 0 g/m^3 in their evaluation. One of the twelve experts consulted by the EPA thought that the shape of the concentration-response function may change at 7 $\mu g/m^3$, suggesting that this level may serve as a possible threshold. A suspected change in the shape of a continuous function by one of 12 experts seems a tenuous basis for saying that any effects below this value should be neglected. As stated in the report, Pope et al. (2002) do show that levels as low as 7 $\mu g/m^3$ can be associated with premature death. Lacking data below that value, that study could not quantitatively assess effects below that value.

The basis for the ascribed threshold seems to be that there is no empirical evidence for mortality effects below the values measured in the ACS study. No evidence other than a single speculation by one of twelve experts consulted by the EPA is provided in support

for the existence of a threshold at all. Applying the proportionality outlined by the proposed methodology to clean regions suggests that the assignment of a threshold may underestimate the premature death rate by 2.5 to 7% for the population in those regions. Lacking some empirical or physiological rationale for assuming that a threshold exists, I seriously question the inclusion of a threshold value.

Question 3: Is the methodology for estimating health impacts avoided by strategies designed to attain the standards reasonable and justified?

The methodology for estimating the health impacts avoided of strategies designed to attain air quality standards is convoluted and confusing. The Ostro reference on which it is supposedly based does not appear in the bibliography, nor does it appear as cited when I do a brief literature search. I have attempted to see if I can rationalize the approach taken. Unfortunately, the meaning or significance of PM_{attain} is not described.

When I go through the algebra for the case where PM_{max} exceeds the standard, I do not recover a meaningful quantity to tell me the meaning or purpose of the reduction factor or PM_{attain}. The statement of the roll-back/attainment model needs to be rewritten to make it clear and unambiguous. It appears that PM_{attain} is intended to mean the PM level that one would estimate from the current year loadings if the PM levels were rolled back to meet the standard.

This would allow for year-to-year fluctuations in PM loadings in estimating health impacts, which seems reasonable.

Given a workable model, existing data would be used to estimate PM concentrations in each census block, using interpolation where local data are not available. Census data would then be used to estimate the population exposed. This seems reasonable. Results from census blocks would then be used to determine population-weighted exposure for each county, and applied to subsequent mortality impact assessments. Since more localized census block assessments are being determined in the methodology, one could also do exposure assessments and mortality impact assessments. Depending upon the nature of the mortality impact model used this could lead to different estimations of mortality than areal averaging of exposure data would suggest.

In estimating the mortality impact, the methodology does not state explicitly what model is to be employed, but rather provides an example of a log-linear function whose origin is not stated. This appears to be the result of applying Poisson statistics to the estimation of the number of deaths occurring in a population. As such, there appears to be a typographical error in the equation which, if I am correct, should read

$$\Delta Y = Y_0 [\exp(\Delta PM) - 1] * pop$$

It should be noted that this model introduces the nonlinearities in the statistics described above that raise questions about the use of county average exposures rather than census tract exposures in estimating mortality effects. Further, its application requires that the mathematical estimation of the change in PM levels be unambiguous, which not the case in the present methodology report.

In summary, the proposed methodology document needs work to make it clear to the

reader. The basis for the proportionality constant is based upon good scientific reasoning. The decision to impose a threshold needs to be better justified if it is to be maintained. Moreover, if it is maintained, the methodology for estimating excess deaths needs to reflect that quantity. The mathematical statements in the report require particular attention to correct a number of apparent errors. The bibliography should include all papers cited.

References

1. Pope CA, Dockery DW. Health effects of fine particulate air pollution: Lines that connect. Journal of the Air and Waste Management Association 56:709-742 (2006).

2. McConnell R, Berhane K, Gilliland F, London SJ, Vora H, Avol E, Gauderman WJ, Margolis HG, Lurmann F, Thomas DC, Peters JM. Air pollution and bronchitic symptoms in Southern California children with asthma. Environmental Health Perspectives 107:757-760 (1999).

3. Gauderman WJ, McConnell R, Gilliland F, London S, Thomas D, Avol E, Vora H, Berhane K, Rappaport EB, Lurmann F, Margolis HG, Peters J. Association between air pollution and lung function growth in southern California children. American Journal of Respiratory and Critical Care Medicine 162:1383-1390 (2000).

4. Gauderman WJ, Gilliland GF, Vora H, Avol E, Stram D, McConnell R, Thomas D, Lurmann F, Margolis HG, Rappaport EB, Berhane K, Peters JM. Association between air pollution and lung function growth in Southern California children – Results from a second cohort. American Journal of Respiratory and Critical Care Medicine 166:76-84 (2002).

5. Pope CA, Burnett RT, Thun MJ, Calle EE, Krewski D, Ito K, Thurston GD. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. Journal of the American Medical Association 287:1132-1141 (2002).

6. Jerrett M, Burnett RT, Ma R, Pope CA, Krewski D, Newbold KB, Thurston G, Shi YL, Finkelstein N, Calle EE, Thun MJ. Spatial analysis of air pollution and mortality in Los Angeles. Epidemiology 16:727-736 (2005).

7. Oberdorster, G. Toxicology of ultrafine particles: in vivo studies, Phil. Trans. Roy. Soc. London A - Math. Phys. Engr. Sci. 358: 2719-2739 (2000).

8. Donaldson, K. and Brown, D. and Clouter, A. and Duffin, R. and MacNee, W. and Renwick, L. and Tran, L. and Stone, V. The pulmonary toxicology of ultrafine particles. J. Aerosol Med. 15: 313-220 (2002).

A.4 Alan Hubbard

UNIVERSITY OF CALIFORNIA, BERKELEY

BERKELEY DAVIS IRVINE LOS ANGELES RIVERSIDE SAN DIEGO SAN FRANCISCO SANTA BARBARA SANTA CRUZ

SCHOOL OF PUBLIC HEALTH
Division of Biostatistics October 1, 2007

Comments on *Methodology for Estimating the Premature Deaths Associated with Long-term Exposures to Fine Airborne Particulate Matter in California.*

Development of a credible range based on expert opinions from a panel of experts selected by U.S. EPA.

This section concerns the standardized methodology used to combine the opinions of 12 experts regarding the health hazards of PM2.5. This results is, per question asked, a set of subjective percentiles characterizing of the probability distribution (sort of an informal posterior probability) of the parameters relating PM2.5 to pre-mature death. For instance, the percentiles of the distribution specifying the slope of the dose-response relationship of PM2.5 and pre-mature death (that is, the change in mortality versus change in 1 ⌠g/m3 of PM2.5). These percentiles characterize both the central tendency of this distribution but also the range of probable values.

I agree that performing a formal aggregation of the expert opinions on the effect-size of PM2.5 exposure as well as providing formal inference would be unwarranted here. First, the sample size is small (only 12) and so any inferential procedures would be based on strong assumptions. Second, it is a stretch to think of this as a random draw of 12 experts from a large population of potential experts, which renders formal inference problematic. So, I think using the median values of the experts' median values seems a reasonable choice for the estimate of the effect size.

Sensitivity Analyses

I am not sure how to interpret taking the upper confidence bound from one study and lower one from other. I think a more defensible method for calculating the uncertainty bounds on the effect estimate would be a more formal method, such as those presented in the sensitivity analyses. For instance, taking the medians of the 95% credible ranges of the various experts. I could also see avoiding the entire expert panel and using the two main studies to derive the estimates and uncertainty bounds. In fact, the sensitivity analyses lead me to think, why not just do a formal meta-analysis since the report appears to be approximating that informally? However, because the analyses do not differ substantially, both in the mean and the range estimates, for the actual estimates and credibility bounds it is a moot point. My only technical comment, which is alluded to in the report, is that two of the studies use the same data and so the analyses formally combining the estimates really only have two independent studies which would certainly

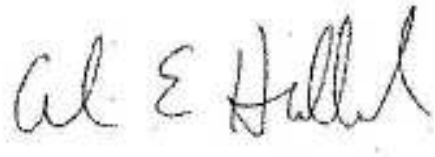
make the confidence limits reported in Table 3 (2 through 5) increase if one accounted properly for this dependence.

Methods for estimating health impacts associated with PM exposures

These appear sensible to me, finding the relative risk for a change in PM2.5 exposure based on the consensus effect size and based on changing each region from it's typical exposure (as described in the report) to the roll-back value (or 0 if roll-back value bigger than typical exposure).

Other Comments

I would add a concern about they main studies that the 12 reviewers did not share, which Jerret, et al. (2005) exemplifies. That is, the adjustment for a large number of confounders in regression models. For instance, Jerret, et al. (2005) adjust for some 40+ confounders. Given how these confounders are entered are typically arbitrary (e.g., linear terms) the final results depend strongly arbitrary choices of model structures. Nonparametric causal inference, assuming you have measured all the confounders, requires that one has an unexposed person precisely matched (on all confounders) for every exposed person. Of course, with continuous exposure and high-dimensional covariates (confounders) this is impossible, so models are assumed. In this case, because the space of possible models is huge, one can only examine a tiny fraction of them, or just arbitrarily choose one. Treating the model as known, which is I know commonly done, really gives distorted inference at the end. The are techniques, which are no panacea, but at least attack this curse of dimensionality in a practical way and provide statistical inference at the end which is more commiserate with the lack of knowledge about the true underlying model. Broadly, these "causal inference" techniques are implemented using inverse weighted procedures (such as estimated of the so-called marginal structure model using inverse probability of treatment weighted estimators) – other more robust estimators are possible. My guess is they would provide at least very different inference (standard errors).



Alan Hubbard
Assistant Professor of Biostatistics
UC Berkeley

A.5 Joel Kaufman

Peer review of draft report entitled "Methodology for Estimating the Premature Deaths Associated with Long-term Exposures to Fine Airborne Particulate Matter in California" draft not dated / version not numbered, but received with cover letter dated August 27, 2007.

Reviewer: Joel Kaufman
General comments:

In general, this is a reasonably well-written description of a methodology, which is basically sound and well-reasoned. I have a few major and a few minor quibbles. I will sort my comments into the sections provided in Attachment II of the mailing, to the extent possible.

1. **Development of a credible range based on expert opinions from a panel of experts selected by U.S. EPA**.

The expert elicitation process seems reasonable as a way to determine a credible range. I am puzzled by the introductory comments which indicate that the process would take into account newer studies, when the expert elicitation did not have access to most of that newer information. In particular, the introduction and Table I include studies not fully considered by the experts in that process. I would advise that the whole process needs to take into account available literature at the time of the document, OR say that you are relying on what was available at the time of the expert elicitation. I think that the dismissal of the Miller et al NEJM paper is a bit facile—since cardiovascular disease is the leading cause of premature mortality and the presumed cause of most PM-related excess mortality, to say that this study can't be included due to not providing estimates of all-cause mortality strikes this reader as difficult to defend. Most epidemiologists strongly prefer research that studies cause-specific mortality to all-cause mortality as being much more robust and meaningful. Again, I would advise that the process either needs to include this study or say that the whole process is based on information published at the time of the expert elicitation. If including the Miller et al paper, I think that the credible range needs to be expanded upward, since this paper not only has a larger magnitude effect-estimate, but also has improved information on exposure measurement, outcome assessment, and control of confounding compared with Six Cities and ACS.

2. **Sensitivity analysis**.

The section on the concentration-response relationship seems reasonable. I presume that the request for peer review is interested in the section on sensitivity analysis included in this section. I think this is basically fine, though I don't think it is reasonable to include both Jerrett and Pope papers in same pooling; should use one or the other. Also, BenMap is not described or cited in full, so a reader doesn't know what this application does "under the hood" and whether it has been validated in some way.

3. Estimation of premature death associated with exposures to PM2.5. Assuming that the issues are resolved with regard to the mortality impact (see comments above), then this seems largely reasonable. I am a bit confused by what was done in Step 4. In particular, does the process take into account the age-distribution for each county? It would seem that age-standardization (between the population in the cohort studies and counties for which projections are being done) would be optimal for this, and if you can't do it for some reason, you need to do some simulations regarding various age-distributions to show that the results are robust to varying age-distributions. I fear that mortality impact forecasting will not be robust to different age-distributions of these counties when compared to the cohorts under study. Step five refers to death rates over the entire population of all ages, then pop refers to population age 30 or above in each county.

The Big Picture

(a) In reading the proposed methodology, are there any additional scientific issues that are part of the scientific basis of the proposed methodology not described above?

No.

(b) Taken as a whole, is the scientific portion of the proposed methodology based upon sound scientific knowledge methods, and practices?

I am mostly concerned about the incorporation or non-incorporation of research published since the expert-elicitation. The methodology needs to be more clear about this.

A.6 Joel Schwartz

Friday, September 28, 2007
Linda Tombras Smith, Ph.D.
Chief, Health and Exposure Assessment Branch
Air Resources Board

Dear Dr Smith

I have reviewed the proposed methodology for the estimation of PM benefits as a result of alternative environmental standards in California. I found the methodology generally reasonable, but felt there was room for improvement. My specific comments are below.

Sincerely,

Joel Schwartz
Professor of Environmental Epidemiology
Harvard School of Public Health
Director, Harvard Center for Risk Analysis

I continue to be puzzled by benefit methodologies that say there is no evidence for a threshold, and then assume a de facto threshold for computing benefits. The only rational way to explain this is that the authors have very strong priors that are virtually immune to data. In that case, the authors owe us:
a) An explanation of those priors
b) A Bayesian analysis that shows us quantitatively how strong the priors had to be to result in the assumption of a threshold.
c) An alternative analysis assuming no threshold.

The likely absence of a threshold means that there would be health benefits associated with reducing exposures even in communities in attainment of the standards. Recognizing this, the European Union has adopted regulations that require percentage rollbacks in all areas, even when in attainment of their guidelines. While it is not my job to recommend alternative regulations, it is worth noting that some approaches to achieving attainment in non-attainment areas will produce reductions in exposure in attainment areas. A good example is the US EPA Clean Air Interstate Rule. These benefits should be estimated, and when a choice of approaches is available to reach attainment, the consideration of those benefits would then be available.

In 1970, Lave and Seskin published a paper regression age standardized mortality rates in US cities against average particle concentrations in those cities. The advantage of that study was that the mortality experience of the entire population of each city was

compared to the average of the population- oriented monitors in the city. While individual exposures differed from the mean exposure, it seemed reasonable to assume that the exposure error was Berkson, and produced no downward bias in the estimated effect, since the average of all persons experience was being compared to the average exposure. The difficulty was that no individual level covariates were controlled, raising questions about confounding (e.g. by SES, smoking, or occupational exposures) and ecological bias.

The studies that EPA and CARB have relied on have alleviated that problem by using cohorts, with individual covariates. The problem with most of those cohort studies is that they are convenience samples, and unlike Lave and Seskin, do not capture the population mortality experience or the population average difference from the monitored exposure. If the convenience sample differed in health and exposure from the population mean identically in all locations, this would be less of a problem. However, there is no reason to believe this is true. Specifically the friends of the ACS volunteers in city A may represent a healthier, and less exposed subset of city A then they do in city B. This, clearly, can introduce bias into the estimates.

First there is potential confounding if, for example, the cities with higher exposures had systematically less healthy subjects recruited. I know of no reason to assume that this bias will always be in the same direction. However, it does introduce a greater uncertainty (above the statistical uncertainty derived from the standard error of the estimate) into the estimate from such a study. Moreover, the greater the possibility of the relation between sample health and population health varying from city to city, the greater this additional uncertainty. Second, there is no longer any reason to assume that the exposure error is predominantly Berkson. This, fairly unambiguously introduces a downward bias.

These concerns apply to all of the cohort studies, with the obvious exception of the Six City Study. The Six City Study chose a neighborhood within each city, recruited a **random** sample of that neighborhood, and put a population oriented monitor in the middle of each neighborhood. Most subjects lived within a few kilometers of that central monitor, and the assumption of Berkson error seems valid. Further, bias due to differential sampling in different locations was eliminated by the random sampling. This means that the extra source of uncertainty, and extra downward bias, present in the other studies is not present in the Six City Analysis, requiring that it be given greater weight. This does not comport with the approach of treating it as the high estimate.

The two studies standing in greatest contrast to this are the Adventist study and the VA study. While the Adventist study recruited from the same population (Adventists) everywhere, they did not necessarily live in locations within counties that had the same relation between exposure and county monitors in each location. While the Methodology discusses this study viz a viz generalizability, this potential source of bias is not discussed. The VA cohort of hypertensives could not control for cardiovascular medicine, despite known large geographic differences in the use of such medicine in hypertensives. For example, beta blockers are more commonly prescribed in the

Northeast than the rest of the US. This presents a substantial risk of confounding, since, for example, sulfate levels are higher than average in the Northeast. In addition the sampling frame is unclear, and may represent a different subset of the population in different cities. Again, the Methodology only discusses generalizability for this study, and not the high potential for bias. Hence I would give these studies less weight, and suggest at least a brief discussion of the issues raised above with respect to all studies.
The second point is that most of the cohort studies, including the original Six City Study, have contrasted a surrogate for long-term exposure with long term survival. They tell us that people live less long in more polluted cities. But the question that CARB needs to answer in order to do an analysis of the benefits of **reducing** air pollution is what mortality reduction accompanies a reduction in exposure. A cross-sectional analysis of mortality rates and air pollution does not tell us that, no matter how sophisticated the Cox proportionate hazard model is. It is an extrapolation to estimate change in mortality for change in pollution. However, the Laden paper provides precisely the estimate that CARB needs. In that sense, it is the only relevant study. Allowing that the extrapolation of the other studies is never the less reasonable, one still needs to give less weight to extrapolations than to studies directly addressing the question. These issues should be recognized and discussed in the health summary. Moreover, the summary of the Laden paper (Table 1) merely quotes the cross-sectional mortality analysis for the extended follow-up, and does not mention, let alone focus on, the coefficient relating change in mortality to change in pollution between two follow-up periods. This should be corrected. Again, greater weight should be given to the Laden study, and it should not be treated as the upper bound estimate.

Regarding the pooling procedure, the methodology correctly identifies issues, such as lack of calibration, which make formal pooling more difficult. However, their central tendency is, in fact, an unweighted median of medians, which is a form of pooling. What is left out is a formal estimate of the statistical uncertainty about that estimate. Instead ranges are taken by looking at the individual studies. That is a reasonable approach, but it could benefit from the alternative, also reasonable approach, of doing a formal estimation of uncertainty.

A meta-analysis has the great advantage of producing an estimate of how much variation among studies is likely due to chance versus true variation in result across study. This could be applied to the underlying studies to estimate statistical uncertainty. Of course, this does not capture the other sources of uncertainty, such as potential confounding, the issues I raised above, etc. That is the reason for expert elicitation—to provide a formal way to capture such uncertainty. That said, the variation in estimates across experts likely reflects both some true variation in how they assess these issue, and interpret the studies underlying their judgment, as well as some stochastic variability. A meta-analysis of their judgments can help estimate how much of the observed heterogeneity across them would be expected by chance and how much represents true uncertainty. Similarly, a Bayesian pooling could examine posterior distributions of estimates based on more or less informative priors. This would be a nice sensitivity analysis to the chosen approach. It would also avoid the difficulty highlighted by the Methodology—that high and low opinions of experts, essentially the outliers of

judgment, would drive the range. The random effects meta-analysis or Bayesian pooling approaches shrink these extremes toward the mean, and provide shrunken range of plausible dose-response curves.

I don't see any problem of pooling Jerrett with Pope, while formally it is a subset of the Pope study, the exposure gradient is entirely within urban area, while Pope's exposure gradient is entirely across urban areas. So these really are different analyses.

I am not sure what Benmap does to estimate random effects meta-analysis. Is it method of moments? Maximum Likelihood? REML? The meta-analysis program in stata will do all three, and I recommend REML.

Inverse distance weighting is a reasonable method for estimating census block level PM2.5 concentrations. If possible some consideration should be given to incorporating traffic density data. For example, regress measured annual PM2.5 at each monitor against traffic density in the block containing the monitor, and use this to adjust the smoothed estimates for each block, which will not otherwise capture the local traffic effects. I recognize this is a nontrivial effort.

Again, I am concerned with the use of a cutoff of 7 $\mu g/m^3$. It not only flies in the face of the expert judgment, it has potentially important consequences. If an strategy to bring one area into attainment results in the lowering of PM2.5 to, for example, 6 $\mu g/m^3$, then CARB will assume there are no health benefits associated with that reduction. Given the empirical and theoretical arguments against a threshold, this would seem to be an approach that would systematically underestimate benefits, and hence systematically bias control strategies towards those that only have local impacts, against those that also impact neighboring locations which are already in attainment. For this reason, I recommend using the background PM2.5 concentration as the cutoff in computing benefits.

I believe that the rollback scenarios are unrealistic. They imply that only locations that exceed the standard rollback by the rollback factor, while sites within the same air basin that meet the standard do not reduce further. But the control strategies that bring the non-attainment sites into attainment will undoubtedly reduce concentrations at all locations in the air basin, regardless of attainment status. Hence this scenario systematically underestimates the benefits of pollution reduction strategies. What if you took the empirical distribution of PM2.5 concentrations in an airshed and rolled the entire distribution down, until the standard was met at all sites. That seems a more likely scenario.

B. Comments on Application to Specific Emission Sources

In this section, a summary of comments from the peer reviewers is presented, followed by individual comments from the six experts.

Summary of Peer Reviewer Comments

On ARB's Proposed Methodology for Estimating Health Impacts Associated with Exposures to Specific Emission Sources

Issue	J. Brook	M Eisner	R Flagan	A Hubbard	J Kaufman	J Schwartz
Modeled Data						
Aggregate grid cells	Yes	Appropriate, but small grid cells may lead to high variability and uncertainty	No comment	Appropriate considering C-R function accuracy	Yes, uncertainties need to be explicitly stated	Yes, appropriate. Errors tend to cancel. Not appropriate to report grid cell result
Applying county incidence rate to smaller area	Within county death rates vary by age, SES.	Yes, but adjust for age and sex distribution of population	Small population samples may introduce systematic uncertainties, in exposure, susceptibility.	Depends on assumptions of C-R function and accuracy of incidence rate.	Yes, appropriate.	Death rates likely higher near port and railyard due to lower SES. Applying county incidence rate underestimates mortality. Age also important
Minimum size population	5,000 to 50,000	Will depend on variability and confidence intervals	No comment	No comment	Depends on confidence intervals.	Pop size determines noise in estimate. Smaller excess death predictions have higher uncertainty
Demographics	Risk will vary by age and health status	Age, sex, race and ethnicity may be different in small pop versus county	Small pop samples may introduce systematic uncertainties, both in exposure, and susceptibility.	If C-R function vary by demographic characteristics, then they become important.	Estimates need to be standardized by age and gender.	Very important. See above.
Single source appropriate	CRF will vary depending on source of PM	Yes, with above caveats	No comment	No comment	Depends on robustness of modeling.	Yes, with concerns above.
Type of source	Yes	Yes	PM from CR function in epi study may differ from the single source. If from DPM, approach may provide lower-	No comment	The method would be applicable and needed in certain regions, esp for ammonium nitrate.	2ndary more uniform and more certain. Wood smoke and traffic likely underestimate mortality because efficient exposure.

Issue		J. Brook	M Eisner	R Flagan	A Hubbard	J Kaufman	J Schwartz
				bound estimate.			
	Other		No comment	No comment	CRF and incidence rate must be same in small/large pops	Emphasize uncertainty at each stage, esp exposure.	No comment.
Emissions Data							
	Appropriate	Only if CRF applies to source and concentration well estimated	Variability and confidence intervals will be an issue	This approach assumes that there is no threshold, which may not be an issue near sources.	Yes, appropriate	Depends on accuracy of emissions inventories.	Yes, appropriate. C-R function may need adjusting. For example, diesel PM may need higher C-R.
	Minimum size	Larger more like CRF	Uncertain	No comment	Same as comments above	Depends on confidence intervals.	Same as comments above
	Demographics	Pop should be like CRF study.	No comment	No comment	Pop demographics should be the same as C-R function.	Estimates need to be standardized by age and gender.	Demographics affect incidence rate.
	Type of source	Yes	Yes, potentially	No comment	No comment	2ndary PM would be more difficult due to chemistry.	Not appropriate to use linear rollback for 2ndary PM because complex chemistry.
	Other	Sensitivity analysis and population mobility;	Is it too imprecise to be meaningful? Is the population exposed to point source similar to epi study population?		Perhaps in log-linear or linear dose-response model, the relative hazard is equivalent to what is proposed, but this will not be true in general.	Emphasize accurate estimates of uncertainty at each stage, esp exposure, and incorporate these uncertainties into calculation of CI.	

B.1 Jeffery Brook

Brief Comments on Proposed Methodology for Estimating Health Impacts Associated with Exposures to Specific Emission Sources

Reviewer: Dr. Jeffrey R. Brook, Senior Research Scientist, Environment Canada; Adjunct Professor, University of Toronto.

The proposed methodologies clearly represent a logical and thought out effort to address the issue of estimating health benefits associated with air pollutant reductions associated with specific sources that tend to impact more localized areas. As in any such assessment, the reliability, representativeness and true meaning of the concentration-response function (CRF) to be applied is an important consideration. The health endpoint(s) considered are also important and a fixation on premature mortality only tells a part of the story. In the applications described in this document the issues of differential susceptibility and differential toxicity/potency of particulate matter of different compositions likely become increasingly important. The former implies that the one CRF may not be ideal, especially in applying it to smaller geographic subsets of the population where there may be spatial clustering of demographic groups (population characteristics) in term of age, race, SES and possible pre-existing conditions that influence susceptibility. Ideally, to get a better feel for these issues, sensitivity analyses based upon a range of realistic assumptions about variability or potential biases, driven by true small scale data on spatial variations in PM levels and composition and population characteristics may provide insights as to how the bottom-line: reduced premature mortality; changes or becomes more uncertain could be helpful.

A. <u>Methodology based on modeled concentrations</u>

The assumes that the incidence rate for the county is the same in each grid cell.
It seems logical to expect that this is variable spatially within a county and areas of higher incidence rate would be pointing towards populations with greater susceptibility and/or greater exposures. SES may be a proxy as could age. It would be worthwhile to examine how these vary among grid squares using census data or any data that might be accessible.
This assumes that the susceptibility distribution of the population in the grid is the same as in the population used to derive the CRF.
There are perhaps two core issues:
Is the CRF the same for different types of PM?
How do we know that all people in an area will see the same size decrease in exposure?
Original CRF's are calibrated to 'area monitors' and so we have some confidence that the changes in mass detected at these monitors reflect the average change in exposure across the population. This is not as safe of an assumption at the local scale.
Responses to Questions for peer reviewers:

Is it appropriate to estimate PM mortality for each grid cell, then sum the results across the grids?
Conceptually, this seems OK, but issues related to the next questions raise concerns.

Is it appropriate to use countywide incidence rates for applications to smaller populations within a given county?
The potential for the validity of this assumption to vary by county seems relatively high. A look at how census-based data on demographics (age, sex, race, SES) varies spatially within counties would shed some light on this. If census data are not sufficient then perhaps property values, percentages of residential property types could be obtained and would be informative. It seems likely there will be counties where variable incidence rate could be expected. Given this, then one needs to consider if the CRF would be different among segments of the population with a higher mortality rate. If we hypothesize that a higher incidence rate is due to a greater prevalence of a pre-existing condition such as TII diabetes (DM) or to an older population then we should expect that a 'general' CRF would be too small. If higher incidence rate is related to SES and the lifestyles that increase the rate then the jury is still out as to whether this itself makes a person more susceptible to air pollution.

How limited can the population size be? What is the minimum affected population size that would make this type of calculation meaningful?
This potentially also varies by geographic region. The more homogenous a population and the more that population is similar to those in ACS and Six Cities the smaller the size that could be considered. Again some sensitivity analyses with census and other spatial data on populations may shed some light on this. If I had to guess I would say 50,000 would mostly likely be safe and there are places were you might be able to get away with about 5000.

Are the population demographics important?
Absolutely, and other variations in susceptibility. There have been acute studies done that show that risk increases with age and it varies depending upon pre-existing conditions such as CHF, DM, COPD, HT, unstable plaque.

Would this methodology be appropriate to estimate the impacts associated with a single source or a limited number of sources of PM?
Most researchers hypothesize that different PM (i.e., from a different source or of a different chemical composition) have different toxicities. Are they different enough to be reflected in a population based CRF is an open question, but it is logical to expect that the CRF should vary by PM type and individual susceptibility. Certainly, evidence grows that traffic PM is a concern and we know about the hazardous nature of DPM. Controlled human exposure studies also seem to get much clearer effects when these are done with diesel exhaust vs. general CAPS from ambient air.

Is the source of PM important in this application?
Yes it is important.

Could this methodology be used if the PM is from gasoline combustion or woodsmoke, or a non-combustion type of source?
If we had robust, population-based CRFs for each source. However, obtaining these and having proof that they are significantly different from one another continues to be very illusive. The issue of endpoints comes up too although the one here is mortality. To some extent, these differences are likely encompassed by the low and high ranges of the CRF and so we may hypothesize that using the same CRF and range (upper and lower bounds) for all different PM includes such variations in the uncertainty or bounds. The issue of what co-pollutants (gases) are associated with the different PM types may magnify the differences between sources, however, and in acute studies (time series) the total risk from two or multi-pollutant models are larger and potentially more stable than just the PM risk alone. See Burnett's et al.'s *Cdn J of Pub Health* paper (Reference below).

Also, in addition to directly emitted PM emissions (primary PM), the conversion of nitrogen oxides to ammonium nitrates (secondary PM) can be modeled. Should one consider the relative contribution of secondary sources compared to the primary PM source in a small population?
Ideally yes, but the PM exposures in the studies that the CRFs have been derived from included both types of PM. One bottom line is that we are getting more confidence that certain PM is more potent, in epi studies, than general PM (e.g., traffic or diesel or possibly metal-enriched PM). Tox studies support this notion of particles being different (DTT assays, etc). But, one should be aware that fine particle nitrate (pNO3), which is semi-volatile does not necessary condense on particles alone. There are likely secondary organics including N-containing species, which can also include amines, that partition more to particles when the thermodynamics also favors pNO3 formation. Of course, actual exposure potential is also an issue here in that in some climates semi-volatile species don't penetrate and/or persist indoors as much as, black carbon, sulfate or heavy PAHs, for example.

What other criteria should be used to determine when such an estimation is appropriate?
Sensitivity analyses

B. <u>Methodology based on emissions data only</u>

Use ARB's estimated county-specific PM2.5 concentrations attributed to diesel sources in year 2000 (CARB 1998).

How well can this be done and what basis is there for assuming that the annual diesel PM concentration is the same across the county when we know it is not? It will be much higher closer to the source(s) of interest, but perhaps one could argue that the actual magnitude varies spatially but the ultimate change in mass will be more uniform across the county. Actually, I don't think so. It might be somewhat more justifiable to assume that the percent change in mass is uniform and then one needs to know the spatial

details of the concentration relative to the population. However, if we consider the typical application of a CRF over a larger area or the actual data used to get the CRF then we have to acknowledge that within that base population there is already

Responses to Questions for peer reviewers:

Is it appropriate to estimate PM mortality based solely on the emissions from a particular source?

This can be reasonable if the relationship between emissions and ambient concentration is linear and correctly quantified (i.e., the data used to get the ratio are reliable and appropriately applied). For the case of DPM, having an appropriate CRF, as opposed to a CRF from total PM2.5, is an issue, just as discussed above.

How limited can the size of the population affected by the emissions from a single source be?
It can be very limited if the source is small, if its location is such that the prevailing winds very consistently blow the emissions in a very consistent direction such as with the sea breeze blowing a plume inland. Other meteorological factors can also limit the size.

What is the minimum affected population size that would make this type of calculation meaningful?
This depends upon the distribution of susceptibility in the population. The larger the size the more likely the distribution will look like the average and more importantly like the population that the original CRF came from. Assuming this is not an issue then the size can be small if the exposure change is known reliably.

What should the population demographics be?
Like those where the CRF came from, in every sense. This was mentioned above.

Is the source of PM important in this application?
Yes, as in my previous discussion.

As described in the previous section, could other sources of PM be considered? It would depend upon the ability to have reliable emissions for the county and reliable estimates of the amount of PM mass in the air that is from that source.

Also, should one consider the relative contribution of secondary PM compared to primary PM?
Same issues as raised above.

What other criteria should be used to determine when such an estimation is appropriate?
Sensitivity analyses of the impact of spatial heterogeneity in population demographics and exposure using reasonable assumptions should be considered.

Population mobility may become more important because the smaller the area influenced by the source(s) of concern the more likely it could be that individuals in the surrounding population move out of the zone of influence regularly and for long periods and also possibly move into the area or at least to where concentrations go up.

References

Burnett R.T., Cakmak S. and Brook J.R., 1998 The effect of the urban ambient air pollution mix on daily mortality rates in Canadian Cities. *Canadian J. Public Health*, **89**(3):152-156.

B.2 Mark Eisner

Proposed Methodology for Estimating Health Impacts Associated with Exposures to Specific Emission Sources

Health impacts from PM exposure are commonly estimated at the statewide or a similarly large geographic scale because (in part) these estimates are based on epidemiologic studies that relied on single ambient air monitoring stations to represent regional exposures to the pollutant. Our interest is in refining and applying such estimation techniques to finer scales, for small populations being affected by small changes in pollutant concentrations that would result from a single or few sources of emissions. The peer reviewers are being asked to comment on these applications.

Below is a summary of two methodologies that could be used to estimate health impacts associated with exposures to PM resulting from specific sources in a limited geographical area. The discussion is divided into two sections based on available information on the pollutant concentration: a) modeled concentrations and b) emissions data. Examples using ports and goods movement are shown to facilitate the discussion.

A. Methodology based on modeled concentrations

In the first scenario, suppose an air dispersion model is used to estimate ambient concentrations of PM in a limited geographic area affected by emissions from a specific source or group of sources. Examples would be locomotive emissions at a rail yard or all sources of diesel (trucks, locomotives, ships) at a California port or harbor. In this scenario, the annual average ambient diesel PM concentration would be estimated by grid cells using a model such as U.S. EPA ISCST3. For each grid cell, the premature death could be estimated based on a concentration-response (C-R) function, the population in that grid cell, and the baseline countywide incidence rates. The total impacts for the affected population in the modeling domain would then be obtained by summing the results from each grid cell.

Hypothetical Example:
A small 2 mile by 2 mile region next to a rail yard within county X has about 10,000 persons over the age of 30, exposed to an estimated diesel PM annual-average concentration of 0.2 $\mu g/m^3$. Using the baseline death rate of 0.009 death/person/year in county X and the C-R function of 10% increase in premature death risk per 10-$\mu g/m^3$ increase in PM2.5 exposure, we would estimate about 0.9 death to result from this small population being exposed to PM.
Questions for peer reviewers:

- Is it appropriate to estimate PM mortality for each grid cell, then sum the results across the grids? It may be appropriate, but one issue will be small grid cell sizes leading to the potential for high variability or uncertainty of results.
- Is it appropriate to use countywide incidence rates for applications to smaller populations within a given county? If this is adjusted for differences in the age and sex distribution of the population, it is probably valid.

- How limited can the population size be? What is the minimum affected population size that would make this type of calculation meaningful? This will depend on variability and the size of the resultant confidence intervals (i.e., if the CI is too wide and the uncertainty is too high, then the results will be less meaningful)
- Are the population demographics important? Yes, because differences in the age-sex distribution of the small population vs. countywide population could introduce confounding. If there are differences in race-ethnicity, effect modification could potentially be an issue as well.
- Would this methodology be appropriate to estimate the impacts associated with a single source or a limited number of sources of PM? Yes, with the limitations above.
- Is the source of PM important in this application? Could this methodology be used if the PM is from gasoline combustion or woodsmoke, or a non-combustion type of source? Also, in addition to directly emitted PM emissions (primary PM), the conversion of nitrogen oxides to ammonium nitrates (secondary PM) can be modeled. Should one consider the relative contribution of secondary sources compared to the primary PM source in a small population? Yes.-
- What other criteria should be used to determine when such an estimation is appropriate?

B. Methodology based on emissions data only

When it is not feasible to model PM concentrations, emissions can be used to estimate health impacts as an alternative methodology. For example, to estimate health impacts associated with Goods Movement activities in the port of Los Angeles, an emissions inventory approach was used as shown below. Details for this methodology can be found in the CARB 2006 report.

5. Use ARB's estimated county-specific PM2.5 concentrations attributed to diesel sources in year 2000 (CARB 1998).
6. Calculate the premature deaths for the base year 2000 by applying a C-R function to the exposed population for a county.
7. Associate the health impacts with the total diesel PM emission inventory for that county in the base year 2000 to determine the number of tons emitted per annual death. This is called the "tons-per-death" factor for the county.
8. Apply the tons-per-death factor to the diesel PM emission inventory for a single source to estimate the average annual deaths associated only with exposure to these emissions, adjusting for population growth between the year of interest and the base year 2000[11]. Note that the diesel PM emissions from the single source may be small compared to the county's emission inventory used in step 3 above.

[11] The impact for year 2005 Goods Movement emissions would be calculated by dividing the emissions by the "tons per death" factor in each county, multiplied by the ratio of year 2005 population over year 2000 population.

Hypothetical Example:

ARB estimated that the diesel PM concentration in county Y is 2 $\mu g/m^3$ for year 2000. This value is used in conjunction with the county's population of 800,000 persons and baseline death rate of 0.009 death/person/year to derive an estimated 136 premature deaths. The total diesel PM emission inventory in county Y is 1,360 tons in year 2000; hence, the tons-per-death factor is 10. A single source which produces 20 annual tons of diesel PM emissions in year 2005 is then estimated to be responsible for about 2.2 premature deaths by using (20 tons/10 tons-per-death) * (880,000 persons/800,000 persons), where 880,000 indicates the county's population in 2005.

Questions for peer reviewers:

- Is it appropriate to estimate PM mortality based solely on the emissions from a particular source? An issue will be the variability of estimates and how wide the confidence interval will be for estimation of the impact of a single source.
- How limited can the size of the population affected by the emissions from a single source be? What is the minimum affected population size that would make this type of calculation meaningful? Uncertain
- What should the population demographics be?
- Is the source of PM important in this application? As described in the previous section, could other sources of PM be considered? Yes, potentially. Also, should one consider the relative contribution of secondary PM compared to primary PM?
- What other criteria should be used to determine when such an estimation is appropriate? Key issues are variability / precision of the estimate (is it too unprecise to be meaningful); genearlizability (is the population exposed to the point source sufficiently similar to that from which health effects estimates were derived); impact of specific types of PM.

References

CARB 1998. California Air Resources Board, Proposed Identification of Diesel Exhaust as a Toxic Air Contaminant. Appendix III. Part A: Exposure Assessment, available at http://www.arb.ca.gov/toxics/id/summary/diesel_a.pdf.

CARB 2006. California Air Resources Board, "Quantification of the Health Impacts and Economic Valuation of Air Pollution From Ports and Goods Movement in California." Appendix A in Emission Reduction Plan for Ports and Goods Movement. March 22, 2006, available at http://www.CARB.ca.gov/planning/gmerp/march21plan/appendix_a.pdf

B.3 Richard Flagan

Comments on Proposed Methodology for Estimating Health Impacts Associated with Exposures to Specific Emission Sources

In my comments on the proposed methodology for county-wide estimation of health impacts, I raised a question that becomes even more important when one seeks to apply that methodology to estimate health impacts of specific emission sources: present

air quality data is obtained using equipment that provides mass based measurements of relatively coarse size fractions (PM2.5 and PM10) from instruments located at a small number of monitoring stations.

Numerous recent studies indicate that small particles that contribute little to the aerosol mass loading may impact health much more significantly than their mass concentrations would suggest (Oberdorster et al., 2000; Donaldson et al., 200). Moreover, studies of health impacts of exposures to ultrafine particles near busy highways (Brunekreef et al., 1997), combined with direct measurements of ultrafine particles as a function of distance from the highway (Zhu et al., 2002; Shi et al., 2001; Zhang et al., 2005; Jacobson et al., 2005), raise questions about the suitability of data obtained at present community monitoring stations for assessing health impacts of some of the sources identified in this proposed methodology; community monitoring stations have traditionally been located some distance from local sources to prevent biasing *samples* in the way that *exposures* will naturally be biased. The aforementioned highway studies reported substantial concentration, and hence exposure, variations over distances of a few hundred meters. The probabilistic health impact model is exponential in PM exposure (if my interpretation of the original methodology report is correct), so averaging exposures over a range with substantial variations will underestimate the health impacts on those individuals closest to the source if such variations are important for the sources of interest. Thus, it is reasonable to consider alternate approaches when addressing individual sources.

Coupling of emission data with an air dispersion model could address these variations in exposure, although the 2 mile by 2 mile grid cell suggested in the example calculation would miss the effects seen in the exposures to diesel emissions near highways carrying significant heavy-duty truck traffic. If the model were based upon the interpolation of data from the carefully sited community monitoring stations, exposure estimates might differ significantly from reality. Moreover, if the model only addresses dispersion, excluding the coagulation, condensation, evaporation, and chemical reaction processes that have been found to lead to the observed rapid variations in fine particle concentrations, exposure estimates may be further compromised.

What is missing in the present epidemiological data is an assessment of the impact of particular constituents of the atmospheric aerosol, either alone or in combination with other constituents of the aerosol or, perhaps, gaseous pollutants. When one applies the broad area results to a specific source, there is a danger that the local emissions doseresponse function may differ dramatically from that of the urban air-shed average. Exposure to a high PM2.5 level near a harbor or rail yard with a large influence of diesel emission would be very different than an equal mass exposure to a marine aerosol (sea salt) at the beach. On the other hand, if the local source is reflective of major pollutant emissions in the urban area, application of the empirical dose-response function could be an excellent approximation; in other cases, it might provide a lower-bound estimate of the health impacts.

The use of the emissions-data-only approach assumes that the health impacts are

strictly linear in particulate-mass concentration; one important implication of this result is that the hypothesized threshold for health effects is dropped from consideration. This may not be an issue since, when near sources, concentrations below the threshold are unlikely. This model assumes that all sources impact health equally on a mass emission basis. Lacking more detailed information about the origins of the health impacts, the approach should provide reasonable estimates.

Small population samples may introduce systematic uncertainties, both in exposure and in susceptibility. On the other hand, applying the health effects correlations to a spatially resolved population exposure may give better estimates of aggregate impacts than would calculations based upon exposures averaged over a city, particularly if the response function is nonlinear.

References

Brunekreef B, Janssen NAH, deHartog J, Harssema H, Knape M, vanVliet P. Air pollution from truck traffic and lung function in children living near motorways. *Epidemiology* 8:298-303 (1997).

Donaldson, K., Stone, V., Gilmour, P.S., Brown, D.M., and MacNee, W. Ultrafine particles: mechanisms of lung injury. Phil. Trans. Roy. Soc. A-Math. Phys. Engr. Sci. 358: 2741--2748 (2000).

Jacobson MZ, Kittelson DB, Watts WF. Enhanced coagulation due to evaporation and its effect on nanoparticle evolution. *Environmental Science and Technology* 39:9486-9492 (2005).

Oberdorster, G. Toxicology of ultrafine particles: in vivo studies. Phil. Trans. Roy. Soc. A-Math. Phys. Engr. Sci., 358: 2719--2739 (2000).

Pope CA, Burnett RT, Thun MJ, Calle EE, Krewski D, Ito K, Thurston GD. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *Journal of the American Medical Association* 287:1132-1141 (2002).

Pope CA, Dockery DW. Health effects of fine particulate air pollution: Lines that connect. *Journal of the Air and Waste Management Association* 56:709-742 (2006).

Shi JP, Evans DE, Khan AA, Harrison RM. Sources and concentration of nanoparticles (< 10 nm diameter) in the urban atmosphere. *Atmospheric Environment* 35:1193-1202 (2001).

Zhang KM, Wexler AS, Niemeier DA, Zhu YF, Hinds WC, Sioutas C. Evolution of particle number distribution near roadways. Part III: Traffic analysis and on-road size resolved particulate emission factors. *Atmospheric Environment* 39:4155-4166 (2005).

Zhu YF, Hinds WC, Kim S, Shen S, Sioutas C. Study of ultrafine particles near a major highway with heavy-duty diesel traffic. *Atmospheric Environment* 36:4323-4335 (2002).

B.4 Alan Hubbard

I will answer the questions below, but make a few general comments first. I am not an expert in risk assessment and I assume this document is addressing different methods of risk assessment given the parameters (dose-response) of PM exposure and baseline population mortality risk have been estimated. My expertise is in estimating these parameters, not risk assessment based on the estimated parameters. Thus, my comments should be taken in this context.

Proposed Methodology for Estimating Health Impacts Associated with Exposures to Specific Emission Sources

Health impacts from PM exposure are commonly estimated at the statewide or a similarly large geographic scale because (in part) these estimates are based on epidemiologic studies that relied on single ambient air monitoring stations to represent regional exposures to the pollutant. Our interest is in refining and applying such estimation techniques to finer scales, for small populations being affected by small changes in pollutant concentrations that would result from a single or few sources of emissions. The peer reviewers are being asked to comment on these applications.

Below is a summary of two methodologies that could be used to estimate health impacts associated with exposures to PM resulting from specific sources in a limited geographical area. The discussion is divided into two sections based on available information on the pollutant concentration: a) modeled concentrations and b) emissions data. Examples using ports and goods movement are shown to facilitate the discussion.

A. Methodology based on modeled concentrations

In the first scenario, suppose an air dispersion model is used to estimate ambient concentrations of PM in a limited geographic area affected by emissions from a specific source or group of sources. Examples would be locomotive emissions at a rail yard or all sources of diesel (trucks, locomotives, ships) at a California port or harbor. In this scenario, the annual average ambient diesel PM concentration would be estimated by grid cells using a model such as U.S. EPA ISCST3. For each grid cell, the premature death could be estimated based on a concentration-response (C-R) function, the population in that grid cell, and the baseline countywide incidence rates. The total impacts for the affected population in the modeling domain would then be obtained by summing the results from each grid cell.

Hypothetical Example:

A small 2 mile by 2 mile region next to a rail yard within county X has about 10,000 persons over the age of 30, exposed to an estimated diesel PM annual-average concentration of 0.2 $\mu g/m^3$. Using the baseline death rate of 0.009 death/person/year in county X and the C-R function of 10% increase in premature death risk per 10-$\mu g/m^3$ increase in PM2.5 exposure, we would estimate about 0.9 death to result from this small population being exposed to PM.

Questions for peer reviewers:

- Is it appropriate to estimate PM mortality for each grid cell, then sum the results across the grids?

Given the accuracy of dose-response model and baseline mortality estimate, I can not see an obvious reasons why this would not be appropriate.

- Is it appropriate to use countywide incidence rates for applications to smaller populations within a given county?

That depends on the modeling assumptions of the dose response and the accuracy of the baseline hazard rate in the small population: is the relative hazard (RH) for a unit increase in PM the same, no matter what the baseline characteristics (is there no effect modification) and can one estimate accurately the baseline hazard in this group?

- How limited can the population size be? What is the minimum affected population size that would make this type of calculation meaningful?

Not qualified to answer this. Depends on how generally the dose-response model applies.

- Are the population demographics important?

In so much as the dose-response model varies by the demographic characteristics, then they become important.

- Would this methodology be appropriate to estimate the impacts associated with a single source or a limited number of sources of PM?

Not qualified to answer this.

- Is the source of PM important in this application? Could this methodology be used if the PM is from gasoline combustion or woodsmoke, or a non-combustion type of source? Also, in addition to directly emitted PM emissions (primary PM), the conversion of nitrogen oxides to ammonium nitrates (secondary PM) can be modeled. Should one consider the relative contribution of secondary sources compared to the primary PM source in a small population?

Not qualified to answer this.

- What other criteria should be used to determine when such an estimation is appropriate?

Just main points above – to determine the number of excess deaths due to PM accurately, requires that the dose-response model and baseline rate, as estimated on a larger population, are the same in smaller sub-populations.

B. <u>Methodology based on emissions data only</u>

When it is not feasible to model PM concentrations, emissions can be used to estimate health impacts as an alternative methodology. For example, to estimate health impacts

associated with Goods Movement activities in the port of Los Angeles, an emissions inventory approach was used as shown below. Details for this methodology can be found in the CARB 2006 report.

9. Use ARB's estimated county-specific PM2.5 concentrations attributed to diesel sources in year 2000 (CARB 1998).
10. Calculate the premature deaths for the base year 2000 by applying a C-R function to the exposed population for a county.
11. Associate the health impacts with the total diesel PM emission inventory for that county in the base year 2000 to determine the number of tons emitted per annual death. This is called the "tons-per-death" factor for the county.
12. Apply the tons-per-death factor to the diesel PM emission inventory for a single source to estimate the average annual deaths associated only with exposure to these emissions, adjusting for population growth between the year of interest and the base year 2000[12]. Note that the diesel PM emissions from the single source may be small compared to the county's emission inventory used in step 3 above.

[12] The impact for year 2005 Goods Movement emissions would be calculated by dividing the emissions by the "tons per death" factor in each county, multiplied by the ratio of year 2005 population over year 2000 population.

Hypothetical Example:

ARB estimated that the diesel PM concentration in county Y is 2 $\mu g/m^3$ for year 2000. This value is used in conjunction with the county's population of 800,000 persons and baseline death rate of 0.009 death/person/year to derive an estimated 136 premature deaths. The total diesel PM emission inventory in county Y is 1,360 tons in year 2000; hence, the tons-per-death factor is 10. A single source which produces 20 annual tons of diesel PM emissions in year 2005 is then estimated to be responsible for about 2.2 premature deaths by using (20 tons/10 tons-per-death) * (880,000 persons/800,000 persons), where 880,000 indicates the county's population in 2005.

Questions for peer reviewers:

- Is it appropriate to estimate PM mortality based solely on the emissions from a particular source?

Seems appropriate to me.

- How limited can the size of the population affected by the emissions from a single source be? What is the minimum affected population size that would make this type of calculation meaningful?

Comments in previous example about the population size apply here.

- What should the population demographics be?

Distribution of demographic characteristics should be the same as those used to estimate the dose-response.

- Is the source of PM important in this application? As described in the previous section, could other sources of PM be considered? Also, should one consider the relative contribution of secondary PM compared to primary PM?

Not qualified to answer this.

- What other criteria should be used to determine when such an estimation is appropriate?

I have a more general comment about the parameter of interest. My guess is that the parameter of interest is the relative hazard (or excess deaths due to PM) comparing the current situation (distribution of PM) to a counterfactual scenario where a particular point source is removed. For instance, using the dose-response model, determine the excess deaths based on current PM concentrations (equivalent in the example to 1360 tons) and a scenario based on the concentration that would result when a particular point source is removed (in the example, 1360-20=1340 tons). Perhaps in a log-linear or linear dose-response model, the results are equivalent to what is proposed, but this will not be true in general.

References

CARB 1998. California Air Resources Board, Proposed Identification of Diesel Exhaust

as a Toxic Air Contaminant. Appendix III. Part A: Exposure Assessment, available at http://www.arb.ca.gov/toxics/id/summary/diesel_a.pdf.

CARB 2006. California Air Resources Board, “Quantification of the Health Impacts and Economic Valuation of Air Pollution From Ports and Goods Movement in California.” Appendix A in Emission Reduction Plan for Ports and Goods Movement. March 22, 2006, available at
http://www.CARB.ca.gov/planning/gmerp/march21plan/appendix_a.pdf

B.5 Joel Kaufman

Proposed Methodology for Estimating Health Impacts Associated with Exposures to Specific Emission Sources

Health impacts from PM exposure are commonly estimated at the statewide or a similarly large geographic scale because (in part) these estimates are based on epidemiologic studies that relied on single ambient air monitoring stations to represent regional exposures to the pollutant. Our interest is in refining and applying such estimation techniques to finer scales, for small populations being affected by small changes in pollutant concentrations that would result from a single or few sources of emissions. The peer reviewers are being asked to comment on these applications.

Below is a summary of two methodologies that could be used to estimate health impacts associated with exposures to PM resulting from specific sources in a limited geographical area. The discussion is divided into two sections based on available information on the pollutant concentration: a) modeled concentrations and b) emissions data. Examples using ports and goods movement are shown to facilitate the discussion.

A. <u>Methodology based on modeled concentrations</u>

In the first scenario, suppose an air dispersion model is used to estimate ambient concentrations of PM in a limited geographic area affected by emissions from a specific source or group of sources. Examples would be locomotive emissions at a rail yard or all sources of diesel (trucks, locomotives, ships) at a California port or harbor. In this scenario, the annual average ambient diesel PM concentration would be estimated by grid cells using a model such as U.S. EPA ISCST3. For each grid cell, the premature death could be estimated based on a concentration-response (C-R) function, the population in that grid cell, and the baseline countywide incidence rates. The total impacts for the affected population in the modeling domain would then be obtained by summing the results from each grid cell.

Hypothetical Example:

A small 2 mile by 2 mile region next to a rail yard within county X has about 10,000 persons over the age of 30, exposed to an estimated diesel PM annual-average concentration of 0.2 $\mu g/m^3$. Using the baseline death rate of 0.009 death/person/year in county X and the C-R function of 10% increase in premature death risk per 10-$\mu g/m^3$ increase in PM2.5 exposure, we would estimate about 0.9 death to result from this small population being exposed to PM.

Questions for peer reviewers:

- Is it appropriate to estimate PM mortality for each grid cell, then sum the results across the grids?
    - Probably it is, but the uncertainties of estimates need to be explicitly stated.

- Is it appropriate to use countywide incidence rates for applications to smaller populations within a given county?
    - This is probably the most appropriate approach
- How limited can the population size be? What is the minimum affected population size that would make this type of calculation meaningful?
    - This needs to be answered based on calculation of the confidence intervals from the calculation.
- Are the population demographics important?
    - Yes, especially age and gender; estimates need to be standardized by age and gender. Race/Ethnicity would be of secondary importance.
- Would this methodology be appropriate to estimate the impacts associated with a single source or a limited number of sources of PM?
    - That depends on the robustness of the source-specific models. I would imagine that a limited number of sources could be modeled robustly in specific areas of the state.
- Is the source of PM important in this application? Could this methodology be used if the PM is from gasoline combustion or woodsmoke, or a non-combustion type of source? Also, in addition to directly emitted PM emissions (primary PM), the conversion of nitrogen oxides to ammonium nitrates (secondary PM) can be modeled. Should one consider the relative contribution of secondary sources compared to the primary PM source in a small population?
    - I would think the method would be applicable and would be needed in certain regions of the state, esp for ammonium nitrate.
- What other criteria should be used to determine when such an estimation is appropriate?
    - I would again emphasize the use of accurate estimates of uncertainty at each stage of the process (especially exposure estimation), and incorporating these uncertainties into the calculation of confidence intervals.

B. <u>Methodology based on emissions data only</u>

When it is not feasible to model PM concentrations, emissions can be used to estimate health impacts as an alternative methodology. For example, to estimate health impacts associated with Goods Movement activities in the port of Los Angeles, an emissions inventory approach was used as shown below. Details for this methodology can be found in the CARB 2006 report.

1. Use ARB's estimated county-specific PM2.5 concentrations attributed to diesel sources in year 2000 (CARB 1998).
2. Calculate the premature deaths for the base year 2000 by applying a C-R function to the exposed population for a county.
3. Associate the health impacts with the total diesel PM emission inventory for that county in the base year 2000 to determine the number of tons emitted per annual death. This is called the "tons-per-death" factor for the county.

4. Apply the tons-per-death factor to the diesel PM emission inventory for a single source to estimate the average annual deaths associated only with exposure to these emissions, adjusting for population growth between the year of interest and the base year 2000[13]. Note that the diesel PM emissions from the single source may be small compared to the county's emission inventory used in step 3 above.

[13] The impact for year 2005 Goods Movement emissions would be calculated by dividing the emissions by the "tons per death" factor in each county, multiplied by the ratio of year 2005 population over year 2000 population.

Hypothetical Example:

ARB estimated that the diesel PM concentration in county Y is 2 $\mu g/m^3$ for year 2000. This value is used in conjunction with the county's population of 800,000 persons and baseline death rate of 0.009 death/person/year to derive an estimated 136 premature deaths. The total diesel PM emission inventory in county Y is 1,360 tons in year 2000; hence, the tons-per-death factor is 10. A single source which produces 20 annual tons of diesel PM emissions in year 2005 is then estimated to be responsible for about 2.2 premature deaths by using (20 tons/10 tons-per-death) * (880,000 persons/800,000 persons), where 880,000 indicates the county's population in 2005.

Questions for peer reviewers:

- Is it appropriate to estimate PM mortality based solely on the emissions from a particular source?
    - The issue again is the accuracy of the emission inventories.

- How limited can the size of the population affected by the emissions from a single source be? What is the minimum affected population size that would make this type of calculation meaningful?
    - This needs to be answered based on calculation of the confidence intervals from the calculation.

- What should the population demographics be?
    - Estimates need to be standardized by age and gender.

- Is the source of PM important in this application? As described in the previous section, could other sources of PM be considered? Also, should one consider the relative contribution of secondary PM compared to primary PM?
    - It is my understanding that estimates of exposure to secondary PM requires modeling as described above so would not be easily done in this approach.

- What other criteria should be used to determine when such an estimation is appropriate?
    - I would again emphasize the use of accurate estimates of uncertainty at each stage of the process (especially exposure estimation), and incorporating these uncertainties into the calculation of confidence intervals.

References

CARB 1998. California Air Resources Board, Proposed Identification of Diesel Exhaust as a Toxic Air Contaminant. Appendix III. Part A: Exposure Assessment, available at http://www.arb.ca.gov/toxics/id/summary/diesel_a.pdf.

CARB 2006. California Air Resources Board, “Quantification of the Health Impacts and Economic Valuation of Air Pollution From Ports and Goods Movement in California.” Appendix A in Emission Reduction Plan for Ports and Goods Movement. March 22, 2006, available at
http://www.CARB.ca.gov/planning/gmerp/march21plan/appendix_a.pdf

B.6 Joel Schwartz

Additional Comments of Joel Schwartz

Here are my responses to the additional questions asked about local risk estimation.

General Comment: The expected value of uncertainty is not zero. Yet if a risk assessment is not done, all statements of qualification to the contrary not withstanding, decision makers will tend to make decisions as if the risk is zero. This is clearly inappropriate. Hence, the appropriate think is always to do a risk assessment, but to appropriately qualify the uncertainties. When the direction of likely bias is known, say that. Equally importantly, when the uncertainty is as likely to be an underestimate as an overestimate, say that. These statements are more important than actual estimates of the magnitude of the uncertainty for three reasons. First, they are subject to less error. It is easier to determine the sign of an effect than its magnitude. Second, they are important for decision making—an intelligent decision maker needs to know if most of the uncertainty would push the estimates in a particular direction. And third, if the estimates are likely unbiased (that is, as likely an overestimate as an underestimate), then while any particular decision may, in the light of future further evidence, have over or under estimated the risk benefit ratio, on average, such decisions will be the correct ones, and that is also important for decision makers to know. So my general comment is yes, do the risk estimation, but….spend a good amount of effort identifying the sources of uncertainty, their likely direction of bias if any, and their likely magnitude. But do something.

One key issue that applies to most of what follows is the question of whether it is best to use the same C-R relationship between PM of different sources and mortality. The most commonly available information is for PM2.5. However, the Dutch cohort, which has just produced a new report, clearly sees effect estimates for traffic particles (measured as BS) that are larger than the average estimate for PM2.5. This suggests that Diesel particles, the major source of black particles, are more than averagely toxic.

Specific Comments:
Questions for peer reviewers:

- *Is it appropriate to estimate PM mortality for each grid cell, then sum the results across the grids?*

  Yes, it is appropriate to estimate mortality within small grid cells, because that better captures the highly non-uniform distribution of the exposure across the county. If done reasonably, and summed across all grid cells within the county, many of the sources of error will tend to cancel out for the sum over the county. It would be appropriate, after paying attention to my further comments below, to report that sum. It would not be appropriate to report the values in each cell, because the high degree of uncertainty within them makes the individual cell estimates too noisy to base more local risk decisions on, and too noisy to communicate to the residents.

- *Is it appropriate to use countywide incidence rates for applications to smaller populations within a given county?*

This requires more care. If the distribution of population characteristics that determine baseline rates is random with respect to exposure, than again, performing such estimates creates random noise, which cancels out when averaged over all cells. However, it is unlikely that this is the case. Consider the examples, such as a port or rail yard, where concentrations of diesel particles likely falls off quickly with distance. Clearly, exposure is concentrated closer to the source. What about susceptibility? Death rates are considerably higher in persons of lower socio-economic condition, and such persons are much more likely to live close to undesirable activities, such as rail yards. Hence it would be appropriate to take this into account. SES data is available on the block group level, which is a geographic area with a typical population of 1500, so this information could be easily obtained by your 4 sq mile grid cells. Baseline mortality rates by county may not be computed by SES routinely. However, it would be better to look at the relative mortality by SES for the entire state, and apply that relative ratio to the County mortality in the county of interest, and then, based on census data, calculate an adjusted baseline mortality rate in each grid cell. Why? Because if poorer people live closer to the sources of emissions and have higher baseline mortality rates, ignoring this is a source of bias, whereas the procedure outlined above has considerable uncertainty, but no obvious bias. No doubt, better approaches could be derived.

A related issue is age. Mortality rates vary considerably by age, and small areas can differ substantially from the county average. I recently did an analysis where an entire census block group was an elderly housing complex. Clearly, it had a considerably different baseline mortality rate. Whether this is an issue or not in your assessments I do not know. But you should certainly check to see whether the population age distribution is different in your 4 square mile cell than in the county as a whole. If they are similar, fine, if there is a substantial difference, you can adjust as above.

- *How limited can the population size be? What is the minimum affected population size that would make this type of calculation meaningful?*

What the population size determines is the noise in the estimate. Meaningful is a different question. A model that predicts 2 excess deaths in one case and 50 in another presumably has considerable uncertainty bands around those estimates. What is less uncertain is that the effect in the first case is smaller than the effect in the second. This is presumably meaningful.

- *Are the population demographics important?*

Very important. See above.

- *Would this methodology be appropriate to estimate the impacts associated with a single source or a limited number of sources of PM?*

Yes, subject to the concerns stated above.

- *Is the source of PM important in this application? Could this methodology be used if the PM is from gasoline combustion or woodsmoke, or a non-combustion type of source? Also, in addition to directly emitted PM emissions (primary PM), the conversion of nitrogen oxides to ammonium nitrates (secondary PM) can be modeled. Should one consider the relative contribution of secondary sources compared to the primary PM source in a small population?*

The approaches are applicable to other sources. Care again, is critical. First of all, since a major source of both bias and uncertainty is the variation in the baseline mortality rate in small areas, sources that are more homogeneously distributed, such as secondary secondary nitrates, are actually easier to deal with. While the attributable risk in each area will be smaller because the risk is not as concentrated geographically, this is actually an advantage in coming up with an estimate of overall effect. Of course, the estimates are only as good as the model, and models for secondary aerosols, whether nitrates, organic carbon, or whatever, have two parts—dispersion, similar to the models for primary particles, and atmospheric chemistry, which adds a layer of complexity and uncertainty. But unless there is evidence from monitored data that the models are biased, it is still reasonable to use them, subject to the usual caveats.

Wood smoke is a tricky one because it is emitted from low chimneys, near where people live, and one expects that the exposure efficiency is much greater than the models would estimate. This is also true for traffic particles, and please examine the literature on exposure efficiency. Basically, the probability of such a particle making it into a lung is greater than would be expected from models that predict exposure on scales of a few hundred meters and larger. There is immediate exposure from the source to people right there that is usually underestimated by models that focus on the pollutant when it is better mixed.

As I noted in the introduction, there is the issue of whether the same PM2.5 coefficient should be used for all sources.

- *What other criteria should be used to determine when such an estimation is appropriate?*

2nd Method: Questions for peer reviewers:

- *Is it appropriate to estimate PM mortality based solely on the emissions from a particular source?*

It is certainly appropriate. The issue is whether the coefficient should be adjusted. I think that the evidence is probably strong enough to suggest a larger coefficient for Diesel particles, and not yet clear for others. After all, what one is presumably estimating is the incremental increase in mortality for an increment in particles. So, if you know enough to use a different C-R, do so. If you don't, then as far as you can tell, the average one (i.e. the one for PM2.5), is appropriate.

- *How limited can the size of the population affected by the emissions from a single source be? What is the minimum affected population size that would make this type of calculation meaningful?*

This is really the same question (and answer as for the other method).

- *What should the population demographics be?*

It is important to take demographics into account as they at minimum affect baseline mortality. There is also some evidence of differential effects of PM.

- *Is the source of PM important in this application? As described in the previous section, could other sources of PM be considered? Also, should one consider the relative contribution of secondary PM compared to primary PM?*

It would not be appropriate to use linear rollback from emissions for secondary particles, as there are substantial nonlinearities in the atmospheric chemistry. Other sources could be considered, but again, as above, some attention needs to be paid to the intake fraction literature.

- *What other criteria should be used to determine when such an estimation is appropriate?*

C. Comments on Diesel PM Methodology

Two of the peer reviewers had expertise relevant to the development of the methodology for estimating diesel PM concentrations. Their comments are included in this section.

C.1 Jeffery Brook

Brief Comments on Methodology for Estimating Ambient Concentrations of Particulate Matter from Diesel-Fueled Engine Emissions

Reviewer: Dr. Jeffrey R. Brook, Senior Research Scientist, Environment Canada; Adjunct Professor, University of Toronto.

In light of the amount of information available, the proposed methodology is reasonable. Linking diesel particulate matter to NO_x concentration is attractive given a relatively large number of NO_x monitoring sites. The cross-checking with the past approach and with an alternate approach to determine α based upon source apportionment helps strengthen the results and ARB has highlighted assumptions and uncertainties and the overall lack of all the desired information clearly in this report.

A few issues to consider (along with the attached marked up version of the report):

The lower DRI estimate is discounted too readily based upon limited arguments. Why were 1995 and 2000 results compared to get 1.8 instead of comparisons this report implies were done for the gasoline-diesel split study? What is the possibility that the lower DRI result is due to changing engine technologies so that diesel particulate matter (DPM) emissions are less relative to NO_x? In comparing these two years and groups the larger number of sample days considered by DRI should also carry some weight in the decision. If the DRI results are reconsidered then this suggests that the source apportionment approach may be leading to a lower α than selected here. This implies that the ARB approach is less conservative.

It should also be noted that the source apportionment approach to get DPM typically apportions the diesel contribution to OC and then scales to total PM. This potentially ignores the amount of EC that is from diesel emissions, as well as some inorganic species such as trace metals and primary sulfate. These additional PM constituents would likely increase the value of α.

In general, given the additional loss mechanism for NO_x in the atmosphere (chemical) compared to fine DPM, the expectation would be that α derived from ambient data would be larger than that derived from the emissions inventory. The results here, using Schauer's source apportionment values, support this. The DRI do not. However, it may well be that the DPM from the emissions inventory does include more than just organics and so the resulting α is larger. Thus, it would be useful for some more information on how the emissions inventory DPM is determined. Is this through the typical applications of the MOBILE emissions model with currently accepted emission factors?

Comments on Methodology for Estimating Ambient Concentrations of

Particulate Matter from Diesel-Fueled Engine Emissions

The proposed methodology would employ NO_x data to estimate particulate matter from diesel-fueled engine emissions. The approach is reasonable given that diesel-fueled engines are responsible for a major portion of the NO_x emissions state-wide, but only a small fraction of the particulate matter. Previously, diesel particulate matter was taken to be proportional to the PM10 mass concentration, a very tenuous assumption even though different scaling factors were applied in urban or rural environments. The use of PM10 is particularly problematic since the mechanisms of formation of particles larger than 2.5μm (or even 1μm) differ dramatically from those that produce smaller particles. The coarse part of the size range of PM10 is dominated by crustal materials; PM2.5 contains less, but still significant crustal and mechanically generated material.

The proposed methodology examines results from Schauer et al., and from the DRI group of Chow and coworkers. One citation is to work of Fujita et al., which appears to be reported only in a web page and is likely to be work that has not undergone critical peer review. The methodology uses the Schauer work as the primary reference. The results of the two studies appear to be in reasonable agreement, at least when corrected by the ratio of the means (1.8). Comparing the correlations shown in Fig. 3 (CHS, 1995) and Fig. 4 (DRI, 2000) one sees striking differences. The earlier study shows a correlation that appears to be consistent with a zero-intercept; the later one has fewer and more scattered data that do not appear consistent with the zero-intercept to which the correlation was forced. One outlier was removed from the early data to improve the fit; the uncertainty in the slope observed when it was not removed was comparable to that obtained in the later data set, a possible indication that the more recent experiments included a broader range of locations than did the earlier ones.

The comparison of the DPM/NO_x ratios suggests that at higher levels the range of values of the ratio decreases, but the means do not vary with the NO_x emission rates. This suggests that the method may provide useful estimates, with some caveats. The NO_x measurements are measured at community monitoring stations. The method proposes using an interpolation method to generate a smooth DPM curve from that sparse data set. For basin-wide exposure estimates, this approach will probably be reasonable; however it will likely underestimate the concentrations near sources because the community monitoring site locations have been chosen to minimize local source effects. Concentrations of some types of diesel particle vary dramatically with distance from highways or other sources as do some health effects (Brunekreef et al., 1997; Zhu et al., 2002; Shi et al., 2001; Zhang et al., 2005; Jacobson et al., 2005). Care will have to be exercised to ensure that the data smoothing does not introduce negative biases in regions that are strongly influenced by local emissions. For basin-wide estimates, this may be a relatively minor point, but it could be important for some calculations.

Minor points on report formatting

The report presents a number of figures, without limited discussion. The figures require

captions that explain what is being plotted; units are also required on the axis labels, e.g., what are the units of DMP/NO_x? One can guess from those plots that do have labels, but the reader shouldn't have to guess. I guess that NO_x refers to tons per day of emissions - again, I shouldn't have to guess.

References

Brunekreef B, Janssen NAH, deHartog J, Harssema H, Knape M, vanVliet P. Air pollution from truck traffic and lung function in children living near motorways. *Epidemiology* 8:298-303 (1997).

Jacobson MZ, Kittelson DB, Watts WF. Enhanced coagulation due to evaporation and its effect on nanoparticle evolution. *Environmental Science and Technology* 39:9486-9492 (2005).

Shi JP, Evans DE, Khan AA, Harrison RM. Sources and concentration of nanoparticles (< 10 nm diameter) in the urban atmosphere. *Atmospheric Environment* 35:1193-1202 (2001).

Zhang KM, Wexler AS, Niemeier DA, Zhu YF, Hinds WC, Sioutas C. Evolution of particle number distribution near roadways. Part III: Traffic analysis and on-road size resolved particulate emission factors. *Atmospheric Environment* 39:4155-4166 (2005).

Zhu YF, Hinds WC, Kim S, Shen S, Sioutas C. Study of ultrafine particles near a major highway with heavy-duty diesel traffic. *Atmospheric Environment* 36:4323-4335 (2002).

C.2 Richard Flagan

Comments on Methodology for Estimating Ambient Concentrations of Particulate Matter from Diesel-Fueled Engine Emissions

Richard C. Flagan, Ph.D.
California Institute of Technology

The proposed methodology would employ NO_x data to estimate particulate matter from diesel-fueled engine emissions. The approach is reasonable given that diesel-fueled engines are responsible for a major portion of the NO_x emissions state-wide, but only a small fraction of the particulate matter. Previously, diesel particulate matter was taken to be proportional to the PM10 mass concentration, a very tenuous assumption even though different scaling factors were applied in urban or rural environments. The use of PM10 is particularly problematic since the mechanisms of formation of particles larger than 2.5μm (or even 1μm) differ dramatically from those that produce smaller particles. The coarse part of the size range of PM10 is dominated by crustal materials; PM2.5 contains less, but still significant crustal and mechanically generated material.

The proposed methodology examines results from Schauer et al., and from the DRI group of Chow and coworkers. One citation is to work of Fujita et al., which appears to be reported only in a web page and is likely to be work that has not undergone critical peer review. The methodology uses the Schauer work as the primary reference. The results of the two studies appear to be in reasonable agreement, at least when corrected by the ratio of the means (1.8). Comparing the correlations shown in Fig. 3 (CHS, 1995) and Fig. 4 (DRI, 2000) one sees striking differences. The earlier study shows a correlation that appears to be consistent with a zero-intercept; the later one has fewer and more scattered data that do not appear consistent with the zero-intercept to which the correlation was forced. One outlier was removed from the early data to improve the fit; the uncertainty in the slope observed when it was not removed was comparable to that obtained in the later data set, a possible indication that the more recent experiments included a broader range of locations than did the earlier ones.

The comparison of the DPM/NO_x ratios suggests that at higher levels the range of values of the ratio decreases, but the means do not vary with the NO_x emission rates. This suggests that the method may provide useful estimates, with some caveats. The NO_x measurements are measured at community monitoring stations. The method proposes using an interpolation method to generate a smooth DPM curve from that sparse data set. For basin-wide exposure estimates, this approach will probably be reasonable; however it will likely underestimate the concentrations near sources because the community monitoring site locations have been chosen to minimize local source effects. Concentrations of some types of diesel particle vary dramatically with distance from highways or other sources as do some health effects (Brunekreef et al., 1997; Zhu et al., 2002; Shi et al., 2001; Zhang et al., 2005; Jacobson et al., 2005). Care will have to be exercised to ensure that the data smoothing does not introduce negative biases in regions that are strongly influenced by local emissions. For basin-wide estimates, this may be a relatively minor point, but it could be important for some

calculations.

Minor points on report formatting

The report presents a number of figures, without limited discussion. The figures require captions that explain what is being plotted; units are also required on the axis labels, e.g., what are the units of DMP/NO_x? One can guess from those plots that do have labels, but the reader shouldn't have to guess. I guess that NO_x refers to tons per day of emissions - again, I shouldn't have to guess.

References

Brunekreef B, Janssen NAH, deHartog J, Harssema H, Knape M, vanVliet P. Air pollution from truck traffic and lung function in children living near motorways. *Epidemiology* 8:298-303 (1997).

Jacobson MZ, Kittelson DB, Watts WF. Enhanced coagulation due to evaporation and its effect on nanoparticle evolution. *Environmental Science and Technology* 39:9486-9492 (2005).

Shi JP, Evans DE, Khan AA, Harrison RM. Sources and concentration of nanoparticles (< 10 nm diameter) in the urban atmosphere. *Atmospheric Environment* 35:1193-1202 (2001).

Zhang KM, Wexler AS, Niemeier DA, Zhu YF, Hinds WC, Sioutas C. Evolution of particle number distribution near roadways. Part III: Traffic analysis and on-road size resolved particulate emission factors. *Atmospheric Environment* 39:4155-4166 (2005).

Zhu YF, Hinds WC, Kim S, Shen S, Sioutas C. Study of ultrafine particles near a major highway with heavy-duty diesel traffic. *Atmospheric Environment* 36:4323-4335 (2002).