State of California

Air Resources Board *Governor Arnold Schwarzenegger*

Mary D. Nichols, Chairman

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

Staff Report

October 24, 2008

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 Exposure to Fine Airborne Particulate Matter in California

Staff Report

Project Coordinator and Lead Author

Hien T. Tran, Ph.D.

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
Melanie Marty, Ph.D., Office of Environmental Health Hazard Assessment

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, Henry Roman and Dr. Katy Walker of Industrial Economics, Inc., for information on the U.S. EPA expert elicitation, Professor Phil Hopke of Clarkson University for his review of the methodology for estimating diesel particulate matter concentrations, and Lindsey Roth of OEHHA for valuable analyses. 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

List of Figures

List of Tables

Acronyms and Abbreviations

ACS	American Cancer Society
AHSMOG	The Adventist Health Study of Smog
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 associated with exposure to airborne 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 associated with PM exposure is based on a peer-reviewed methodology developed by the U.S. Environmental Protection Agency (EPA) for their health 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 impact analyses. This report discusses the results of staff's review of the recent scientific literature related to the mortality effects associated with exposure to fine PM (PM2.5) and presents revisions to the current methodology.

In this report, the relative risk of premature death associated with PM2.5 exposure was reevaluated based on a review of all relevant scientific literature, and a new relative risk factor was developed. This new factor is a 10% increase in risk of premature death per 10 $\mu g/m^3$ increase in exposure to PM2.5 concentrations (uncertainty interval: 3% to 20%). Using this new factor, staff estimates that in the year 2005, diesel PM contributes to 3,500 premature deaths statewide (uncertainty interval 1,000 to 6,400). Also, staff estimates that exposure to ambient PM2.5 concentrations above 5 $\mu g/m^3$ can be associated with about 18,000 premature deaths statewide annually, with uncertainty ranging from 5,600 to 32,000 deaths, based on 2004-2006 air quality data.

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 an independent 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. In addition, all public comments received on the May 22, 2008 draft version of the report have been incorporated into this staff report. Specific responses to individual comments are addressed in Appendix 5.

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 associated with public exposure 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 2004a, 2004b, 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 associated with reducing population exposure to ambient PM that would result through the implementation of the proposed measures (CARB 2003a, 2003b, 2003c, 2004a, 2004b, 2004c). However, such quantified health impacts assessments are not required in a regulation. They are only used as an indicator of the types of health benefits that would likely accrue due to implementation of a proposed regulation.

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 number of premature deaths. In U.S. EPA's health impacts 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 exposure 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 impact analyses. These recent studies prompted CARB staff to consider updating the PM2.5 mortality relationship. For example, Jerrett et al. (2005a) 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 exposure show that the PM-mortality relationship can be larger than predicted by daily time-series studies (Samet et al. 2000a, 2000b). 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. As summarized by Schwartz et al. (2008), the potential mechanisms for the association include: changes in autonomic function, perhaps leading to increased risk of arrhythmias, changes in inflammation and thrombotic factors, potentially increasing the risk of myocardial infractions, impaired endothelial function, and exacerbation of respiratory illness.

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 health impact 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 exposure 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' exposure 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 exposure 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), with an added discussion on publications that appeared after June 2006.

Daily time-series studies of acute exposure 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.

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

A.1 Original Harvard Six Cities and ACS Studies

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.

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, exposure) 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

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

Table 1: Summary of key studies and relative risks on long-term exposure to PM and premature death (adapted from Pope and Dockery 2006, with results published after June 2006 added)

			Percent Increases in Relative Risk of Mortality (95% CI)		
Study	**Primary Source**	**Exposure Increment**	**All Cause**	**Cardio-pulmonary**	**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)
Harvard Six Cities, extended analysis, linearity explored	Schwartz et al. 2008	10 µg/m^3 PM2.5	10 (0, 21)†		
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. 2005a	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. 2006a	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, extended analysis	Beelen et al. 2008	10 µg/m^3 BS		22 (-1, 50)[c]	3 (-12, 20)
Netherlands, extended analysis	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)	–	–
Medicare Cohort in ACS locations	Eftim et al. 2008	10 µg/m^3 PM2.5	10.9 (9, 12.8)	–	–
Medicare Cohort in Harvard Six Cities Study locations	Eftim et al. 2008	10 µg/m^3 PM2.5	20.8 (14.8, 27.1)	–	–
Medicare Cohort in eastern U.S.	Zeger et al. 2008	10 µg/m^3 PM2.5	6.8 (4.9, 8.7)	–	–
Medicare Cohort in central U.S.	Zeger et al. 2008	10 µg/m^3 PM2.5	13.2 (9.5, 16.9)	–	–
Medicare Cohort in western U.S.	Zeger et al. 2008	10 µg/m^3 PM2.5	-1.1 (-3.0, 0.8)	–	–

†Based on current year's exposure only and may be an underestimate for long-term impacts; [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.

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.

Also, on page 197 of the HEI Reanalysis report, Figure 21 represents a cross-tabulation of "residual" (or excess) mortality risk based on 1982-1989 follow-up (after controlling for 44 individual covariates such as smoking, diet, etc) graphed against PM2.5 concentration recorded between 1979 and 1983 using the Inhalable Particle Network dataset (50 cities). The residual mortality for Los Angeles is moderate compared to the other 49 cities, and so are the PM2.5 concentrations, suggesting that the information from Los Angeles is consistent with the rest of the cohort. The figure clearly shows excess risk (above zero risk) for Los Angeles.

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 exposure, 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 exposure, 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 exposure, 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 partially reversible over periods of possibly as short as a year.” 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.

Subsequent to the peer review of this report, Schwartz et al. (2008) examined the linearity of the concentration-response function of PM2.5-mortality using data from the extended follow-up to the Harvard Six Cities Study. The results show that the response function is in agreement with Laden et al. (2006). Further, it was found that it is linear down to background levels.

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 exposure 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 exposure 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 exposure. 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 exposure 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 exposure 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), generally agrees with 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, their results lend convincing support to the general 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 exposure were estimated by linking fixed-site ambient monitoring data with resident zip codes. 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 due to respiratory causes. A statistically insignificant increase was observed for death from all causes.

Medicare Cohort Air Pollution Studies

After this report underwent the formal peer review process, several new publications appeared in the literature. Eftim et al. (2008) recently published a long-term PM2.5 exposure study based on a cohort retrospectively developed from the Medicare database. The cohort and PM2.5 monitoring data were from the same cities and counties as those included in the Harvard Six Cities Study (SCS: Dockery et al., 1993; Laden et al., 2006) and the ACS Cancer Prevention Study (CPS) II cohort (ACS: Pope et al., 1995; 2002; 2006) studies. The cohort included over 7.3 million people who were likely more broadly representative of the American population than the SCS and ACS cohorts in terms of demographics, including race, income, and education. However, the database has several limitations, in that it only includes people over 65 years of age, and there is no information on potential confounders and effect modifiers, such as smoking or body weight. Eftim et al. adjusted for these factors, and several socioeconomic factors on a county level based on census data. The results for the period 2000-2002 are higher than reported previously in the ACS and SCS publications.

Zeger et al. (2008) expanded on Eftim et al's. (2008) study to include over 13 million Medicare enrollees and a five year exposure assessment (2000 – 2005). The study employed methods similar to Eftim et al. (2008), although it also included secondary analyses on a regional basis (Eastern, Central and Western United States), and based on age (65 to 74, 75 to 84, and over 85 years of age). The results showed a decrease in effect with advancing age, such that there was no effect in persons over 85 years of age, consistent with the results of Enstrom (2005) and Laden et al. (2006).

Several factors could influence the differences between Eftim's and Zeger's results and those from the ACS and SCS studies. The size of the exposure aggregation units in larger metropolitan areas could lead to underestimated effect estimates. The Medicare cohort was generally older than the SCS and ACS cohorts, and several papers have suggested that the influence of PM2.5 on mortality decreases in the oldest age groups (Enstrom, 2005; Laden et al., 2006). However, effect estimates for the Medicare cohort may be biased upward due to lack of adjustment for individual level risk factors. It is possible that there are other confounding factors that were not corrected for, although the Health Effects Institute reanalysis of the SCS and ACS studies (Krewski et al., 2000) showed that adjustment for smoking and other individual characteristics had little influence on the effect estimates.

In Zeger et al. (2008), the statistically significant results for eastern and central United States are in general agreement with previous publications. However, Zeger et al. (2008) found no significant effect of PM2.5 on mortality in the western United States. This result may be due to lack of control for individual level covariates in the analysis. These covariates may include body mass index, diet, lifestyle, or other factors that differentiate the Los Angeles basin from other counties in the West. Further, "the West" was defined as urban areas of California, Oregon, and Washington. Thus, in the stratified analysis, the authors effectively compared the Los Angeles basin with other parts of the region. The authors recognize that the Los Angeles basin counties have higher PM levels than other West Coast urban centers, but lower adjusted mortality

rates. Therefore, the result for the West can be significantly biased by the lack of control for individual-level lifestyle factors.

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 exposure, 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 associated with air pollution is a difficult task. As recommended by both the National Research Council (2002) and the Science Advisory Board (U.S EPA 2004b), 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 exposure to PM2.5. This process asked the 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 exposure can contribute to the cardiovascular disease process, thus adding to the plausibility of the positive association between exposure and disease found in the long-term cohort studies.

B. U.S. EPA elicitation process

In this section, we quote extensively from a report by U.S. EPA's contractor, Industrial Economics (IEc, 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 uncertainty be performed for health impacts analyses, including estimating premature death associated with exposure to PM2.5 levels.[2] As a result, U.S. EPA conducted an expert elicitation study with twelve experts to better characterize the uncertainty in the estimated reductions 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 the results from the U.S. EPA-sponsored expert elicitation study. Each expert in the elicitation study considered relevant theoretical and empirical evidence available at the time of the study. Experts' were encouraged to consider evidence that both supported and cast doubt on a PM2.5-morality relationship.[3]

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

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

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 (IEc, 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

- 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 (IEc, 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 exposure 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 exposure 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 $\mu g/m^3$ decrease in ambient annual average PM2.5 may reflect reductions in both short-term peak and long-term average exposure 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

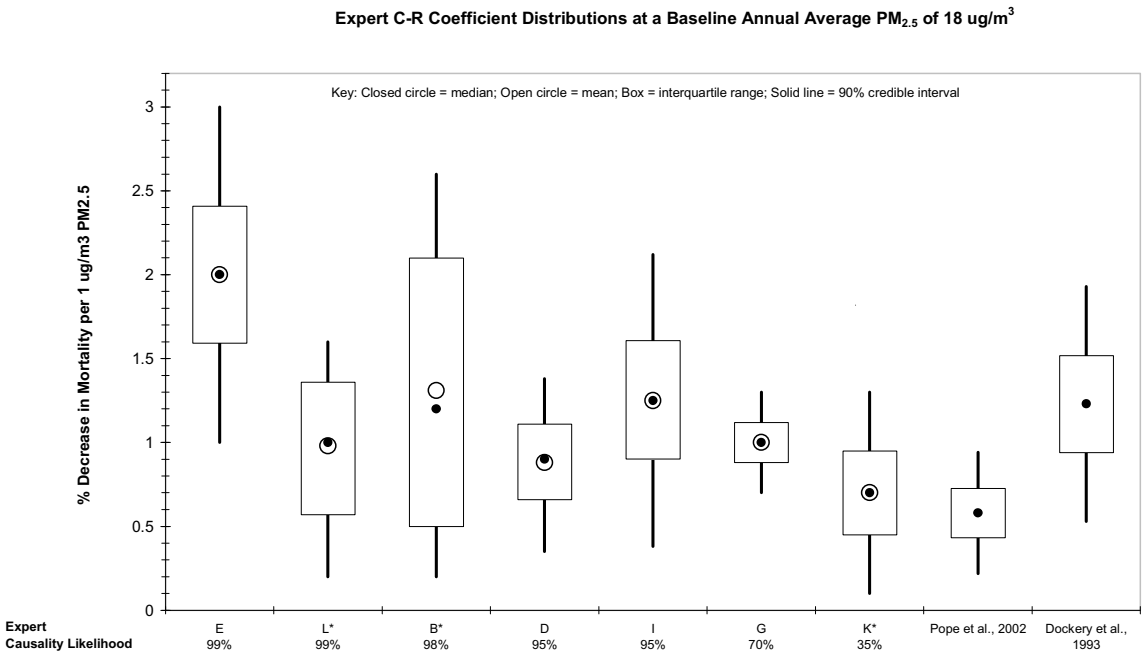
Figures 1 and 2 display the experts' responses to the main quantitative elicitation question. The distributions provided by each expert, identified by the letters A through L, are depicted as box and whisker plots with a solid circle symbol showing the median (50th percentile); an open circle showing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

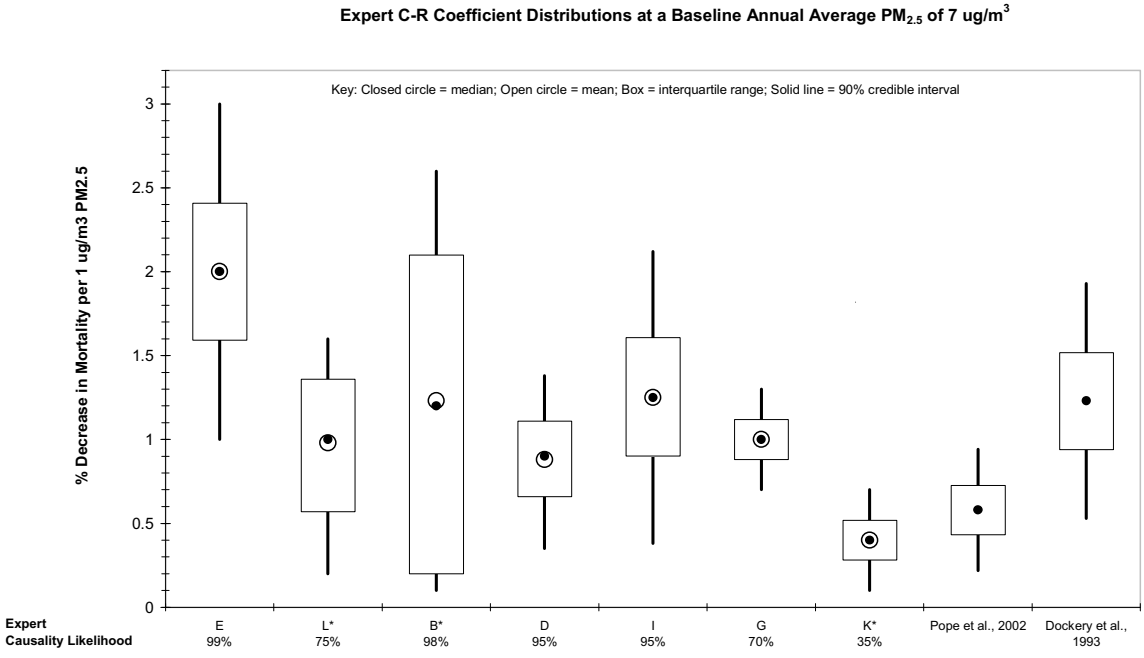
Each expert's stated best estimate of the likelihood of a causal relationship between PM2.5 and premature death is shown on the x-axis and the experts are arrayed in order of decreasing certainty of causality. Figure 1 displays the distributions for the experts who chose to provide a distribution conditional 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 $\mu g/m^3$ and 7 $\mu g/m^3$, 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 $\mu g/m^3$ decrease in annual average PM2.5 exposure. 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 average annual exposure to PM2.5.

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





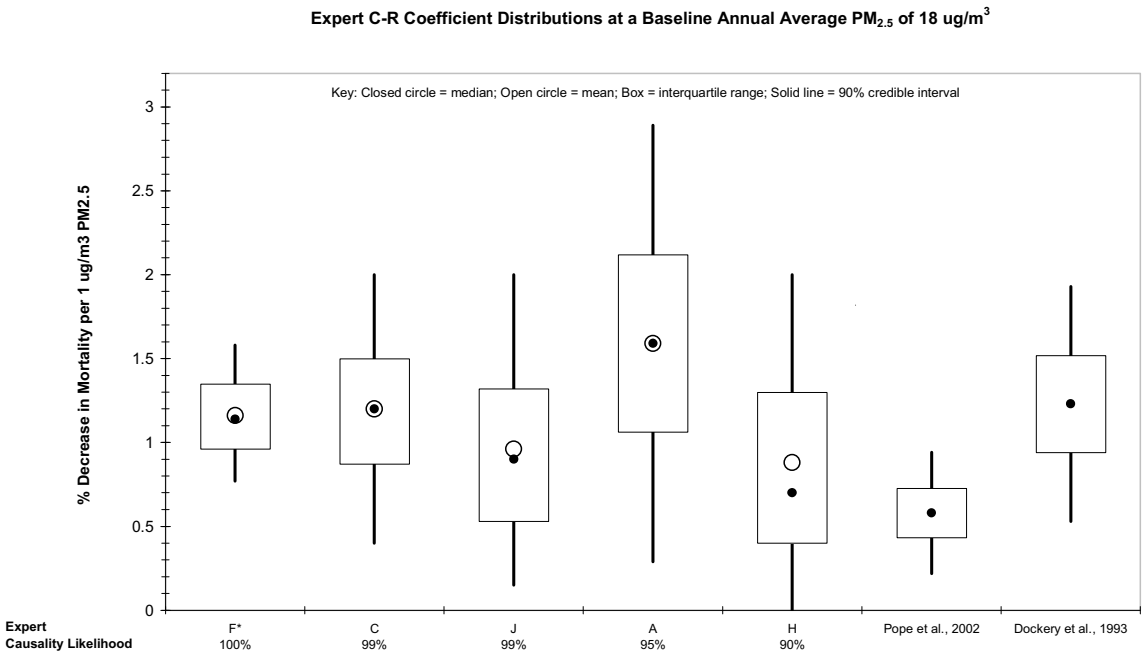
Note: Box plots represent distributions as provided by the experts to the elicitation team. When asked, experts in this group preferred to give conditional distributions and keep their probabilistic judgment about the likelihood of a causal or non-causal relationship separate.

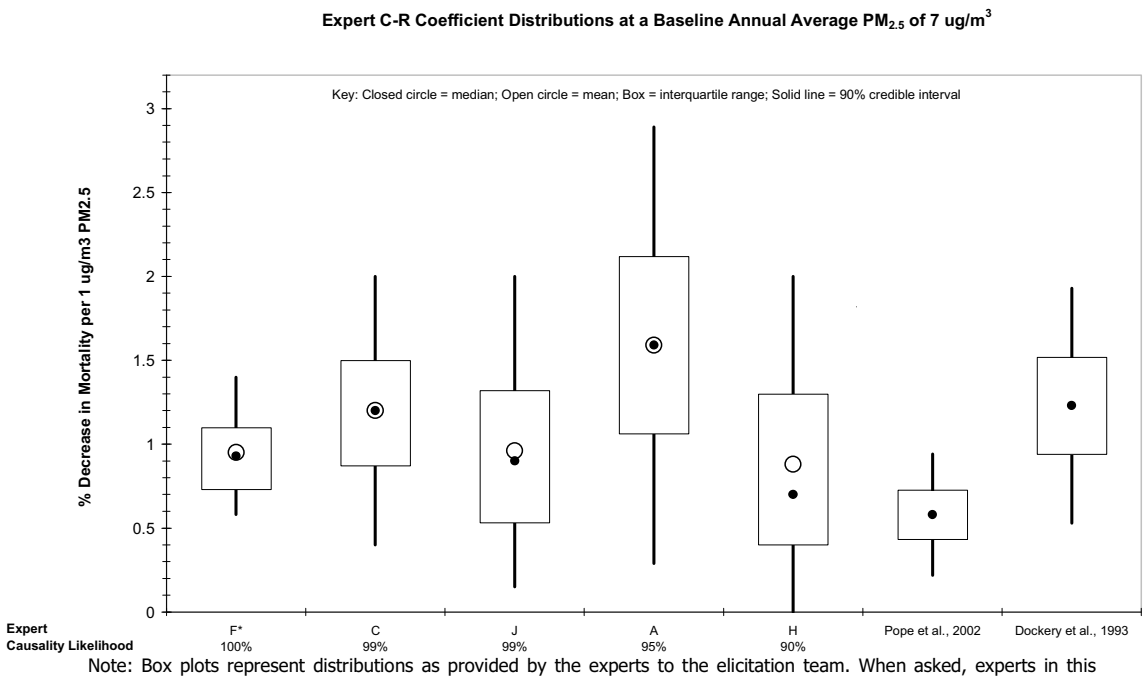
*Experts' C-R coefficient distribution changes between 7 and 18 μg/m^3.

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 μg/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 (IEc, 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. 2005a, 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. 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 $\mu g/m^3$. 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 was modest.
- The 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). 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 exposure, all experts based their median effect estimates on effects due to long-term exposure. Short-term exposure effects were sometimes used to derive lower-bound effect estimates.
- The 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.
- The 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. 2005a; 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 (IEc, 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 (IEc, 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 & 2005a	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 (IEc, 2006).

	ACS (Pope et al., 2002)	ACS LA Reanalysis (Jerrett et al., 2005a)	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 Environment
  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 health impacts 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 the general population for use in health impacts analyses. While the relative risk in premature death per unit change in PM2.5 long-term exposure 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 appropriately interpret the

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

expert 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 such as the Washington University-EPRI Veterans Cohort (Lipfert et al. 2000, 2003, 2006a, 2006b), 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, and Enstrom (2005) study focused on the elderly population only, as do Eftim et al. (2008) and Zeger et al. (2008). In addition, studies on infant mortality (Woodruff et al. 1997, 2006) do not directly address long-term exposure 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. 2002, 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 exposure 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. (2005a) 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 to develop a relative risk for use in quantifying public health impacts should have controlled for 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. Although the Enstrom (2005) study of elderly Californians assigned PM2.5 on the basis of data from just a few monitoring sites and at times on very few measurements (Brunekreef 2006), the studies by Pope (1995, 2002) used the same exposure data set. Accurately matching county of residence and exposure to fine PM is a second concern. While 90% of the cohort lived in the same county in 1983 as 1972 (Enstrom 2006), on average only about 66% of the cohort lived in the same county in 1999 as in 1972 (Enstrom 2005). In contrast, as discussed in section II.A, it is noteworthy that the data from the original ACS study by Pope (1995) and Harvard Six Cities study by Dockery (1993) were independently verified and re-analyzed by Krewski et al. (2000); the findings confirmed the validity of the previously published findings.

Another issue of importance in longitudinal studies is the time of follow-up since the initial enrollment of a cohort. The most significant reason for not giving greater weight to the Enstrom (2005) publication is the 40 year follow-up. At first glance, this long follow-up is an attractive idea. However, the Cox proportional hazards model is influenced by long-term trends that make it unlikely that the assumption of proportionality to the hazard is valid for a period of 40 years. Some examples of such trends would be changes in health care, and the relationship between aging and death (Janes et al. 2007). While it is unlikely that changes in health care, land use, demographics and other risk factors will vary significantly on the scale of a few years, they will, and in some cases have, changed considerably over 40 years. The influence of these long-term

trends on the risk of mortality associated with prolonged exposure to PM2.5 is not addressed in Enstrom (2005). The original ACS and Six Cities studies were less than ten years in duration, reducing the likelihood that this issue applies to them. However, as follow-up in these cohorts continues, this will increasingly become an issue, unless updates to model adjustments for these factors are made.

It is also likely that at some point across a 40 year period the risk of dying in any given year dwarfs any additional risk added by PM2.5, making additional risk related to PM2.5 undetectable in an older cohort, as is likely the case in Enstrom (2005). As the subjects move into this age category, it will become very difficult to distinguish additional risk due to PM2.5 exposure from that related to aging. Such is suggested by Zeger et al. (2008) as well. In fact, Enstrom (2005) demonstrates this, in that the relative risk for a PM2.5 effect on death decreases through the latter measurement periods reported in the paper. It should be noted that Enstrom's relative risk for the 1973 to 1983 time period is similar to that reported in the first analysis by Pope et al. (1995) using the same exposure data, and when the subjects in the two cohorts were of similar age.

Exposure characterization is an important aspect of the analysis of longitudinal data analysis. As was done in Pope et al. (2002) and Enstrom (2005), a single estimate of exposure to PM2.5 was developed for each metropolitan region or county. Later analysis by Jerrett (2005a) of the ACS cohort in Los Angeles shows that as exposure estimation is refined with sophisticated modeling, the effect increases (RR 1.17 compared to RR 1.06 in Pope et al.), suggesting that exposure classification can strongly influence the association between PM2.5 and increased risk.

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. 2005a). 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. (2005a) 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 health impacts 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. 2005a) 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. (2005a) produced results similar to those found in the Harvard Six Cities studies. Thus, it is justifiable to use the 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 exposure 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 likely not the case. Some expert elicitation practitioners use a series of calibration exercises, utilizing questions for which the answer is known or knowable, to assess 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. They were obtained by applying a California PM2.5 data set to the twelve expert distributions in BenMAP, taking into account the

distributions that are conditional on a causal relationship and the distributions that vary with PM2.5 concentrations. Consequently, the credible range can be treated as a 90% uncertainty interval around the estimate of the PM-mortality relationship.

D.1 Sensitivity Analysis

To determine the robustness of the proposed methodology for developing the concentration-response relationship, various methods for pooling the twelve experts' distributions can be used. 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. (2005a) 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. (2005a) using inverse-variance weighting — to give more weight to studies with tighter confidence bounds than those with wider confidence bounds.
3. Pooling Pope et al. (2002) with Dockery et al. (1993), using equal weight.
4. Pooling Pope et al. (2002) with Dockery et al. (1993), using inverse-variance weighting.

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

5. Pooling Pope et al. (2002), Laden et al. (2006), and Jerrett et al. (2005a) using random effects in BenMAP[5].
6. Pooling Pope et al. (2002) and Laden et al. (2006) using random effects in BenMAP.
7. Pooling all 12 expert distributions using random effects in BenMAP.
8. 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.

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

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

[6]http://www.epa.gov/ttn/ecas/regdata/RIAs/

detailed discussion of this weighting scheme.

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

E. Methodology for estimating health impacts associated with PM exposure

In this section, we discuss the methodology developed to estimate the health impacts associated with PM exposure above a predetermined level. This methodology is consistent with that used in a CARB 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 concentration (not three-year average concentrations) is 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.

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

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

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 estimate no PM-related mortality (a 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 tract level. After each change in health impacts is calculated for each census tract, 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 CARB's consideration of a cut-off level. It is defined as the lowest level above which PM-related mortality can be quantified, or the level below which we estimate no PM-related mortality. Recent evidence suggests that exposure 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 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., 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

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

study population would experience an increased risk of death). From a theoretical and conceptual standpoint, all experts generally thought 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 the best means of addressing the population threshold issue, suggesting this approach is best for evaluating 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 high variation in exposure levels and a long follow-up period. The following is a discussion of three alternatives for a cut-off level level: 7, 2.5, and 5 $\mu g/m^3$, and staff recommendation.

Cut-Off Level of 7 $\mu g/m^3$. The level of 7 $\mu g/m^3$ is 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 exposure to PM2.5 levels as low as 7 $\mu g/m^3$ can be associated with premature death. Since the ACS study is the largest cohort study of a general population conducted to date, it would be reasonable to use 7 $\mu g/m^3$ as a cut-off level for calculating PM2.5-related mortality. However, there is limited direct empirical evidence that some effects are likely to occur at lower levels, as discussed below. Since the evidence for long-term mortality effects at levels below 7 $\mu g/m^3$ is limited, quantifying human health impacts associated with exposure to lower levels requires personal judgment and inference from the available data on long-term studies.

Cut-Off Level of 2.5 $\mu g/m^3$. One possibility is to select the background level for PM2.5 as the cut-off level, which addresses all impacts associated with exposure to anthropogenic PM. In California, the background PM2.5 level is 2.5 $\mu g/m^3$ (Motallebi et al. 2003). 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, because the study population consisted of older women only, we could not justify using these results for a general population. Subsequent to the peer review of our report, Schwartz et al (2008) provided evidence for the linearity of the C-R function close to background levels.

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 exposure. The reviewers recognized that selecting a cut-off level involves professional judgment due to limited empirical evidence in the low PM2.5 concentration 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). Subsequent to the peer review of our report, Eftim et al. (2008) found significant relationships between long-term exposure to PM2.5 and mortality at levels as low as 6 $\mu g/m^3$ using 2000-2002 data.

Staff Recommendation. While empirical evidence indicates that mortality can be associated with long-term exposure to PM2.5 levels as low as 6 $\mu g/m^3$, the consensus of the peer reviewers is that effects are likely to occur down to the level of 4 to 5 $\mu g/m^3$. Therefore, in consideration of the more recently published reports and the outcome of our independent peer review, staff recommends that the cut-off level be 5 $\mu g/m^3$.

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 CARB 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 evaluation risk to small populations exposed to PM2.5 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 (Pope et al. 2002, Laden et al. 2006), and incidence rates are obtained 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 exposure to PM resulting from specific sources in a limited geographical area. The methodologies are based on the information available on the pollutant concentration: a) modeled concentrations and b) emissions data. Information on ports and goods movement is shown as examples. As staff develops these methodologies more fully, they will be made available for peer review and public comment.

G.1. Methodology based on modeled concentrations

In this 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 at grid cells using a model such as CALPUFF. 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.

H. Peer review process

Following the Cal/EPA External Scientific Peer Review Guidelines, CARB staff submitted a formal request to the Cal/EPA Project Director for the review of an earlier draft of the report. In it, staff clearly listed the scientific issues relevant to the proposed methodologies in the staff report and stated the required expertise in the reviewers for a successful evaluation of the proposed methodology. Also, staff submitted a listing of individuals who may have a conflict of interest, including our scientific advisors and the experts in the U.S. EPA expert elicitation. Reviewer candidates were independently identified by the University of California at Berkeley, Institute of the Environment, in collaboration with UC colleagues. Each candidate was required to complete a Conflict of Interest Disclosure form, which was reviewed by the Cal/EPA Project Director for the independent peer review. Candidates were accepted as reviewers only if the disclosure information showed they had no conflict of interest related to the report.

Six reviewers were identified by UC Berkeley and selected by the Cal/EPA Project Director to review the proposed methodology. 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 evaluated whether CARB staff correctly interpreted the results published in the literature, including U.S. EPA's expert elicitation, and whether staff has correctly developed methods for estimating premature deaths associated with public exposure to ambient PM. Following a 30-day period, the peer reviewers provided staff with written comments on an earlier draft of the report. Staff then addressed and incorporated the results of this peer review into a draft report for public release on May 22, 2008. In this report, the final methodology has taken into account all public comments received by July 11, 2008. Details on the process and the results of this peer review can be found in Appendix 5.

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 exposure, with a 3 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 exposure 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 exposure, 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 $\mu g/m^3$ change in PM2.5 exposure, 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 $\mu g/m^3$ increase in PM2.5 exposure

Scenario	*Low*	*Mean*	*High*
Proposed credible range	3%	10%	20%
1. three studies, equal weight	2%	12%	26%
2. three studies, inverse-variance weight	4%	11%	19%
3. two studies, equal weight	2%	15%	30%
4. two studies, inverse-variance weight	1%	14%	34%
5. three studies, random effects pooling	3%	11%	19%
6. two studies, random effects pooling	3%	10%	20%
7. twelve experts, random effects pooling	0%*	10%	21%
8. ten 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 exposure to ambient PM

In this section, we present the results of estimating premature deaths associated with ambient PM exposure above certain cut-off levels. Tables 4a shows the number of premature deaths using a 10 percent relative risk associated with exposure to PM2.5 above 5 $\mu g/m^3$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 extrapolated 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-4c is discussed in Section IV.

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

Air Basin	*Low*	*Mean*	*High*
Great Basin Valleys	1	3	6
Lake County	1	1	5
Lake Tahoe	<1	1	2
Mojave Desert	62	200	360
Mountain Counties	35	110	210
North Central Coast	27	90	160
North Coast	20	68	120
Northeast Plateau	2	8	13
Sacramento Valley	370	1,200	2,200
Salton Sea	56	190	330
San Diego County	370	1,200	2,200
San Francisco Bay	800	2,600	4,700
San Joaquin Valley	740	2,400	4,200
South Central Coast	130	440	780
South Coast	3,000	9,700	17,000
Statewide Total	**5,600**	**18,000**	**32,000**

*Totals do not add up due to rounding. Air quality data from years 2004 to 2006.
Although the population-weighted concentrations for Lake Tahoe Northeast Plateau Air Basins (shown in Appendix 1) are below 5 µg/m^3, the estimated deaths in this table are non-zero due to some census tracts having concentrations higher than 5 µg/m^3.

As shown in Tables 4a above, exposure to 2004-2006 PM2.5 can be associated with approximately 18,000 premature deaths statewide annually, with an uncertainty ranging from 5,600 to 32,000 deaths. For comparison, if a cut-off level of 7 µg/m^3 were used, see me about this about 14,000 premature deaths (uncertainty range: 4,300 to 25,000) would be estimated to be associated with PM2.5 exposure. Also, if health effects were assumed to occur down to non-anthropogenic background of 2.5 µg/m^3, approximately 24,000 premature deaths (uncertainty range: 7,200 to 41,000) would be estimated annually.

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

In addition to examining the mortality impacts associated with exposure above certain PM2.5 levels, we also assessed the health benefits of attaining the established ambient air quality standards. Tables 4b and 4c 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 $\mu g/m^3$ and the State standard of 12 $\mu g/m^3$, respectively. For this calculation, the cut-off level was set at 5 $\mu g/m^3$. Details on the methodology used in calculating these estimates are provided in Appendix 2.

Table 4b: Annual premature deaths avoided by attainment of the national annual PM2.5 standard of 15 $\mu g/m^3$ *

Air Basin	*Low*	*Mean*	*High*
Great Basin Valleys	<1	<1	<1
Lake County	<1	<1	<1
Lake Tahoe	<1	<1	<1
Mojave Desert	8	26	47
Mountain Counties	1	4	7
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**	**9,800**

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

Table 4c: 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	39	71
Mountain Counties	7	24	44
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	700	1,300
San Joaquin Valley	450	1,500	2,700
South Central Coast	13	45	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 exposure to diesel PM exposure

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 exposure, with a 3 to 20 percent confidence interval. The concentration of ambient diesel PM was calculated using ambient NO_X concentrations. A full discussion of the methodology for estimating diesel PM from NO_X concentrations can be found in Appendix 3.

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

Air Basin	*Low*	*Mean*	*High*
Great Basin Valleys	<1	<1	<1
Lake County	2	8	14
Lake Tahoe	<1	<1	<1
Mojave Desert	19	65	120
Mountain Counties	8	26	47
North Central Coast	6	20	37
North Coast	4	12	22
Northeast Plateau	<1	<1	<1
Sacramento Valley	53	180	320
Salton Sea	10	35	64
San Diego County	83	280	510
San Francisco Bay	160	540	980
San Joaquin Valley	75	250	460
South Central Coast	20	68	120
South Coast	610	2,000	3,700
Statewide Total	**1,000**	**3,500**	**6,400**

*Year 2005, based on the new PM2.5-mortality relative risk of 10 percent per 10 μg/m^3 increase in PM2.5 exposure. Totals do not add up due to rounding.

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

In this section, results are presented based on the application of the two methodologies discussed in section II.G.

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 II.G.1) in the area near the Ports of Los Angeles and Long Beach. Using the new PM2.5-mortality function of 10 percent per 10 μg/m^3 change in PM2.5 exposure, staff estimated that based on modeled diesel PM concentrations for year 2002, the annual premature deaths associated with the ports' emissions are approximately 120, with uncertainty interval of 36 to 310 deaths. The population data from the 2000 Census was extrapolated 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 II.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 activities statewide.

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

Pollutant	*Low*	*Mean*	*High*
Primary Diesel PM	630	2,000	3,900
Secondary Diesel PM (Nitrates)	500	1,600	3,100
Secondary Diesel PM (Organic Aerosols)	15	49	95
Other Primary PM2.5[2]	12	39	75
Statewide Total[3]	***1,200***	***3,700***	***7,100***

[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 exposure; 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 exposure to PM2.5 and the risk for premature death.

Up to now, CARB staff has calculated mortality impacts associated with PM2.5 exposure based on the C-R relationship from the American Cancer Society study (Krewski et al. 2000, Pope et al. 2002). Several recently published studies prompted CARB staff 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 an uncertainty interval from 3 to 20 percent). Although the interpretation of the recent literature mostly favors a no-threshold model, staff discussed several possible cut-off levels and recommended using a cut-off of 5 μg/m^3 for health impacts quantification. As shown in Tables 4a above, exposure to 2004-2006 PM2.5 can be associated with approximately 18,000 premature deaths statewide annually, with an uncertainty ranging from 5,600 to 32,000 deaths. For comparison, if a cut-off level of 7 μg/m^3 were used, about 14,000 premature deaths (uncertainty range: 4,300 to 25,000) would be estimated to be associated with PM2.5 exposure. Also, if health effects were assumed to occur down to non-anthropogenic background of 2.5 μg/m^3, approximately 24,000 premature deaths (uncertainty range: 7,200 to 41,000) would be estimated annually.

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 exposure to diesel PM in year 2005 can be associated with about 3,500 deaths, with uncertainty ranging from 1,000 to 6,400.

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 exposure, 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. (2005a), 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 that approximately 120 premature deaths (uncertainty interval: 36 to 310) are associated with annual PM2.5 exposure to emissions resulting from port 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 exposure 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.

Concentration-Response Function

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, time-series and national cohort 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. 2005a; 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 each study's 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 of 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 of the health impacts tables, which represent 95 percent confidence intervals.

Baseline Mortality Rate

Baseline mortality rates were entered into the C-R functions in order to calculate the estimates presented in this report. 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 used the same baseline rates as the U.S. EPA. Additional information was obtained from the 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 as toxic as PM2.5. However, this approach may underestimate the true effects of diesel PM exposure on adverse health outcomes. Indirect evidence for this possibility comes from a number of studies that link motor vehicle-related PM exposure to premature death including:

- Elderly people living near major roads had almost twice the risk of dying from cardiopulmonary causes (Hoek et al., 2000).
- PM from motor vehicles was linked to increased mortality (Tsai et al., 2000).
- Fine PM (PM2.5) from mobile sources accounted for three times the mortality as did PM2.5 from coal combustion sources (Laden et al., 2000).

There is also some direct evidence for responses of human subjects specifically exposed to diesel PM. In two inhalation studies, the researchers examined lung inflammatory response after controlled exposure to diesel PM (Nightingale et al, 2000; Salvi et al., 1999). In one study, Nightingale et al. exposed healthy volunteers to 200 $\mu g/m^3$ diesel PM for 2 hours at rest. The researchers found increased inflammatory markers in sputum in exposed individuals compared to air exposed volunteers. The study suggests that exposure to diesel PM in high ambient concentrations leads to airway inflammatory response in healthy volunteers. The other study exposed healthy volunteers to 300 $\mu g/m^3$ diesel PM (Salvi et al, 1999). They found at these levels, diesel PM exposure produces systemic and pulmonary inflammatory response as measured in airway lavage and endobronchial biopsies.

There is also evidence of immunotoxicity associated with diesel PM (for example Diaz-Sanchez et al., 1997, 1999). Diesel PM exposure was found to increase the allergic response in individuals who had a positive skin test to ragweed compared to ragweed alone (Diaz-Sanchez et al., 1997), and nasal challenge with diesel PM induced sensitization to a new allergen in atopic individuals (Diaz-Sanchez et al., 1999).

None of the available epidemiologic studies of PM has measured the diesel content of the outdoor pollution mix. However, the extensive animal toxicology literature on the health impacts of constituents of diesel exhaust PM leads to the conclusion that diesel exhaust PM is at least as toxic as the general ambient PM mixture. Since CARB staff has made quantitative estimates of the public health impacts associated with diesel exhaust PM exposure based on the assumption of equal toxicity, the estimates may underestimate the true 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 CARB 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 CARB 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 exposure was developed: 10 percent increase in premature death per 10 $\mu g/m^3$ increase in PM2.5 exposure (uncertainty interval: 3 percent to 20 percent). Also, staff assumed a cut-off level of 5 $\mu g/m^3$ as the lowest level of PM2.5 that is associated with a change in risk for premature death based on the latest published literature and the peer reviewers' recommendation. Using this approach, staff estimates that exposure to ambient PM2.5 can be associated with about 18,000 premature deaths statewide annually, with uncertainty ranging from 5,600 to 32,000 deaths, based on 2004-06 air quality data. 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.

Brignell, J. Sorry, Wrong Number! Brignell Associates (2000)

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 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.

Diaz-Sanchez, D, M P Garcia, M Wang, M Jyrala, A Saxon. Nasal Challenge with Diesel Exhaust Particles Can Induce Sensitization to a Neoallergen in the Human Mucosa. J Allergy Clin Immunol. (1999) Dec ;104 (6):1183-8 10588999.

Diaz-Sanchez, D, A Tsien, J Fleming and A Saxon. Combined diesel exhaust particulate and ragweed allergen challenge markedly enhances human in vivo nasal ragweed-specific IgE and skews cytokine production to a T helper cell 2-type pattern. Journal of Immunology, Vol 158, Issue 5 2406-2413, (1997).

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.

Eftim, S.E.; Samet, J.M.; Janes, H.; McDermott, A.; Dominici, F. Fine Particulate Matter and Mortality, A Comparison of the Six Cities and American Cancer Society Cohorts with a Medicare Cohort; *Epidemiology* (2008), 19:209-216.

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

Enstrom, J.E. Response to “A Critique of 'Fine Particulate Air Pollution and Total Mortality Among Elderly Californians, 1973-2002” by Bert Brunekreef, PhD, and Gerard Hoek, Ph', *Inhalation Toxicology* (2006), 18:7, 509 — 514

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.

IARC (International Agency for Research on Cancer) Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Humans: Diesel and Gasoline Engine Exhausts and Some Nitroarenes, Vol. 46, IARC, World Health Organization, Lyon, France. 1989.

IEc, 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.

IEc, 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.

Janes, H.; Dominici, F.; Zeger, S.L. Trends in Air Pollution and Mortality: An Approach to the Assessment of Unmeasured Confounding. *Epidemiol.* (2007), 18:416-423.

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).

Krewski, D.; Burnett, R.T.; Goldberg, M.S.; Hoover, K.; Siemiatycki, J.; Abrahamowicz, M.; White, W.H. Validation of the Harvard Six Cities Study of Particulate Air Pollution and Mortality. New England J. Medicine. 350, 198-199 (2004).

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., 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.

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)

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).

Nightingale, J.A.; R. Maggs; P. Cullinan; L.E. Donnelly; D.F. Rogers; R. Kinnersley; K.F. Chung; P.J. Barnes; M. Ashmore; A. Newman-Taylor. Airway Inflammation after Controlled Exposure to Diesel Exhaust Particulates. Am. J. Respir. Crit. Care Med., Volume 162, Number 1, July 2000, 161-166

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.; 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., 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.; 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; 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,; 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.

Salvi, S.; A. Blomberg; B. Rudell; F. Kelly; T. Sandström; S.T. Holgate; A. Frew. Acute Inflammatory Responses in the Airways and Peripheral Blood After Short-Term Exposure to Diesel Exhaust in Healthy Human Volunteers. Am. J. Respir. Crit. Care Med., Volume 159, Number 3, March 1999, 702-709.

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.

Schwartz, J.; Coull, B.; Laden, F.; Ryan, L. The Effect of Dose and Timing of Dose on the Association between Airborne Particles and Survival. Env Health Persp (2008) 116, 1: 64-69.

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.

Tsai FC, Daisey JM, Apte MG. An Exploratory analysis of the relationship between mortality and the chemical composition of airborne particulate matter. Inhalation Toxicology 12 (Supplement 2): 121-135, 2000.

US EPA (2002) Health Assessment Document for Diesel Engine Exhaust. National Center for Environmental Assessment, Office of Research and Development, U.S. Environmental Protection Agency, Washington, D.C.

U.S. Environmental Protection Agency. (2004a). 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. (2004b). 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.

Zeger, S.L.; Dominici, F.; McDermott, A.; Samet, J.M. Mortality in the Medicare Population and Chronic Exposure to Fine Particulate Air Pollution in Urban Centers (2000-2005). *Environ. Health Perspect.* (2008), online August.

APPENDICES

Appendix 1 (PM2.5 Exposure)

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**

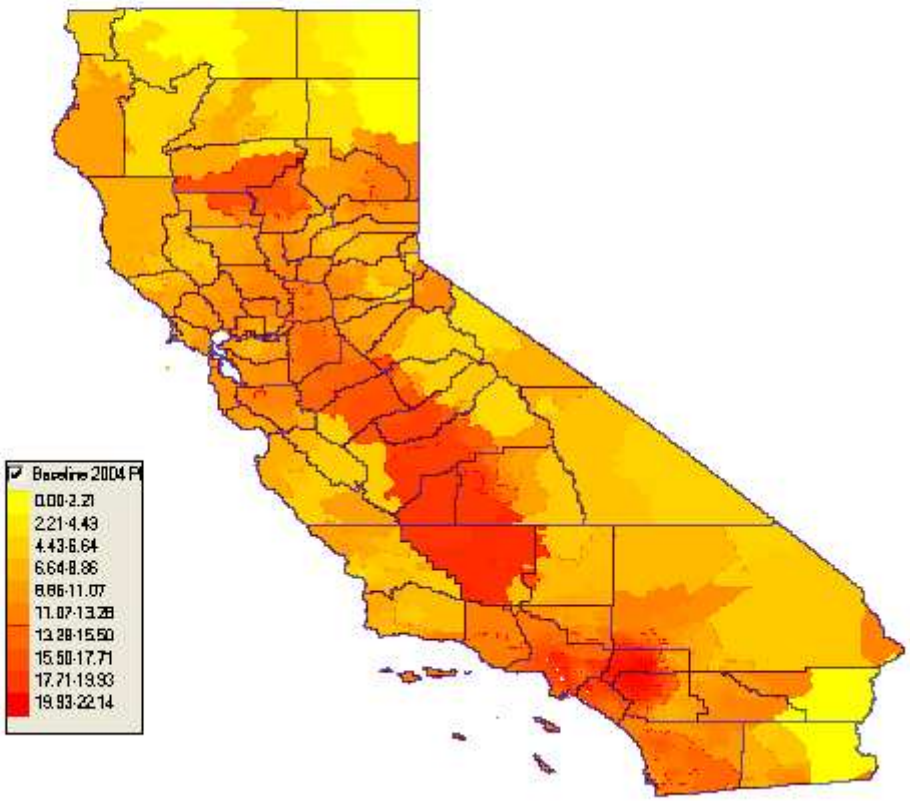
PM2.5 Air Quality Monitoring Program in California

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, CARB 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:

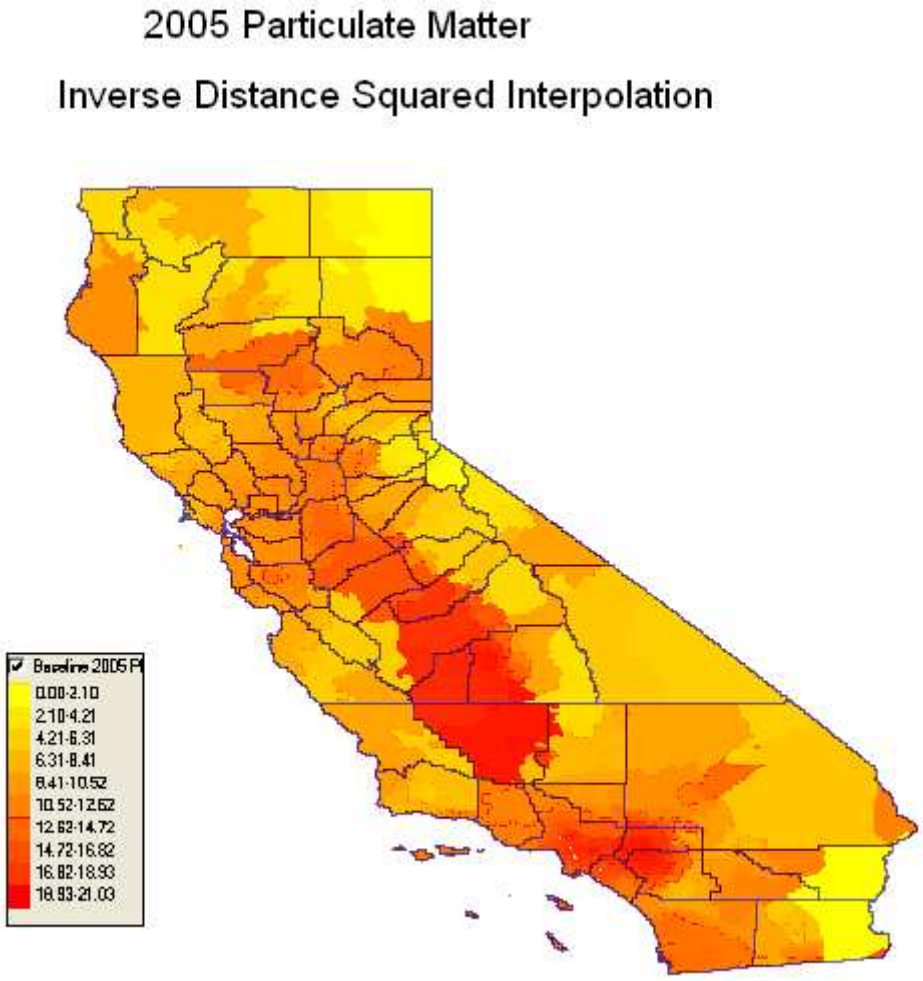
- 2004-2006 PM2.5 mass from the Federal Reference Method (FRM) monitors. California's PM2.5 monitoring network now includes 90 FRM monitoring sites. The FRM sites collect 24-hour mass data using federally approved methods, which means they satisfy specific federal regulatory requirements.
- 2004-2006 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



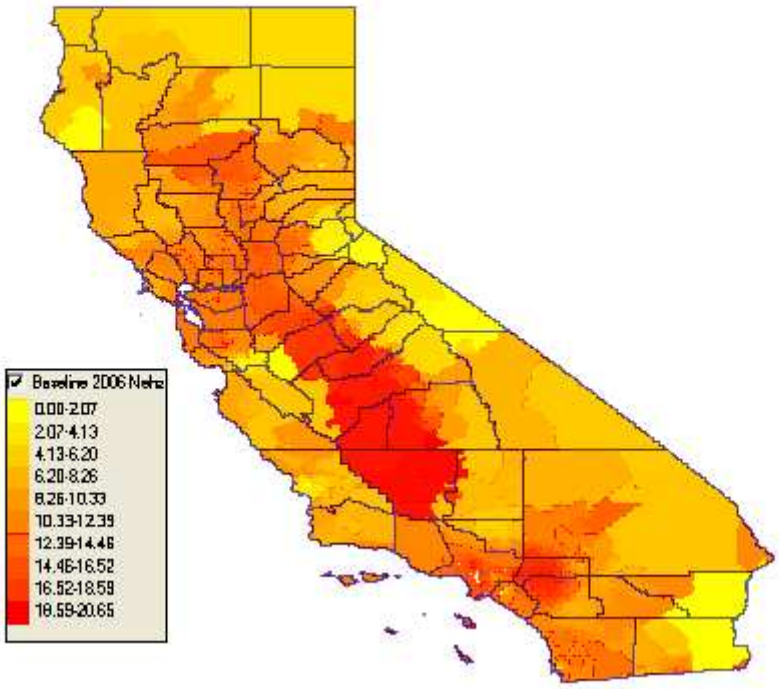
Population weighted PM2.5 concentration for the year 2004 interpolated using BenMAP. The values in the key are in units of $\mu g/m^3$.



Population weighted PM2.5 concentration for the year 2005 interpolated using BenMAP. The values in the key are in units of $\mu g/m^3$.

2006 Particulate Matter

Inverse Distance Squared Interpolation



Population weighted PM2.5 concentration for the year 2006 interpolated using BenMAP. The values in the key are in units of $\mu g/m^3$.

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 exposure 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 exposure 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:

$$S = \alpha C, \qquad (1)$$

where α is a scale factor that is independent of geographical 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 (NO_x is not a unique tracer for diesel emissions). 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 ambient 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, because of the 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. In addition, NO_x emissions in California are primarily from mobile sources, with diesel vehicle emissions accounting for approximately half the on-road mobile contribution and for greater than 80% of the off-road contribution. The limited variation of the diesel engine emissions contribution to total NO_x is a reflection of the similarity of on- and off-road fleet composition and activity in different air basins. In this respect, California likely differs from other regions of the country in the scarcity of important point stationary sources (such as power plants and refineries). Contributions from such point sources would introduce proximity dependencies and preclude the use of a simple NO_x-scaling methodology to approximate DPM.

Methods

In this section, we develop an approximate value for α, the ratio of ambient DPM to total 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 for the year 2005.

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.

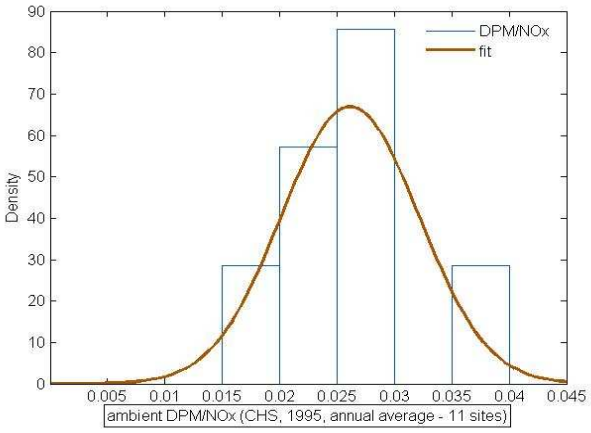


Figure 1

Results

Source apportionment of PM collected in the South Coast Air Basin was done by J. Schauer as part of the Children's Health Study (CHS) in 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. Although each filter PM sample was collected over a two week interval, filters from each site were composited into three seasonal time periods. Each composited sample was analyzed for organic marker compounds and utilized in chemical mass balance source apportionment modeling. We concentrate on using annual

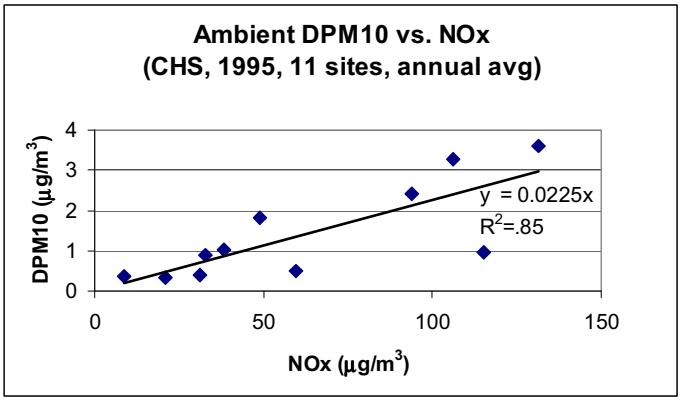


Figure 2

average results from the apportionment study, and show seasonal trends below. 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). 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 influential value (for Mira Loma) gives a slope of 0.026 (0.006), which is also shown in Fig. 2.

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 individual measurements for the ratio DPM/ NO_x retain site-specific variability in concentrations, we believe the associated statistics are better estimates than regression coefficients for DPM exposure-related work. We take the standard deviation of the distribution of DPM/ NO_x values as the measure of uncertainty in α for SA studies.

Figure 3 shows plots of the CHS data for each of the three composited time periods. The slope exhibits a clear seasonal dependence with largest value in summer and smallest in winter. This variation can not be explained completely by EI estimates (summer, winter), which show much smaller variation, and indicates further sources of uncertainty in the use of short timescales for scaling NO_x.

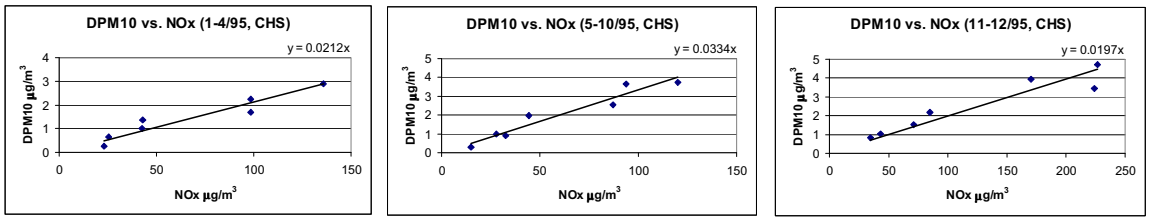


Figure 3

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 4 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.

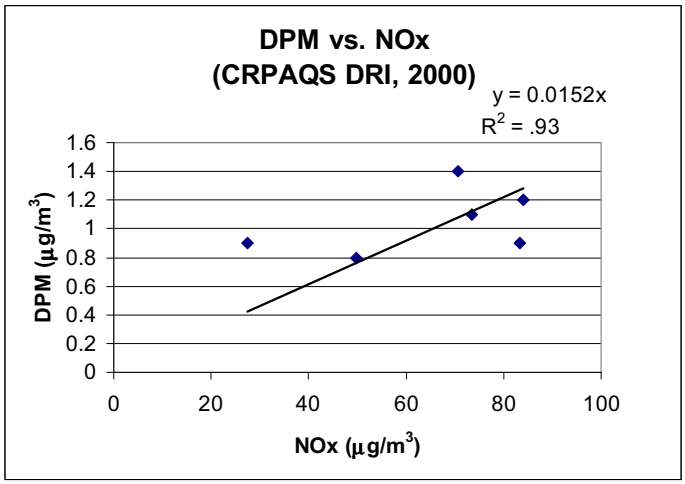
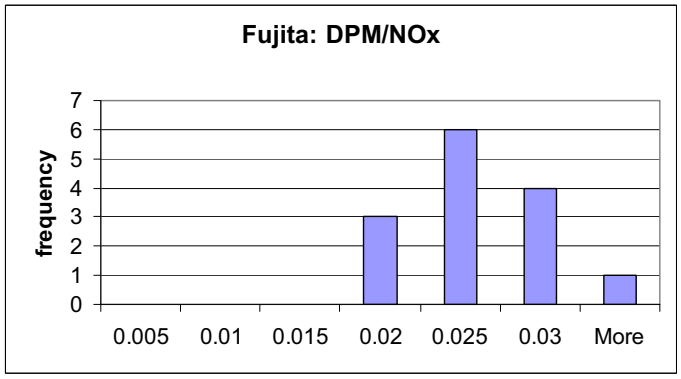


Figure 4

We note that the relative variability of DPM/ NO_x in both studies is very large: standard deviation/average ≈ .5 (.011/.024, .009/.017), commonly referred to as the coefficient of variation. 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.

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)



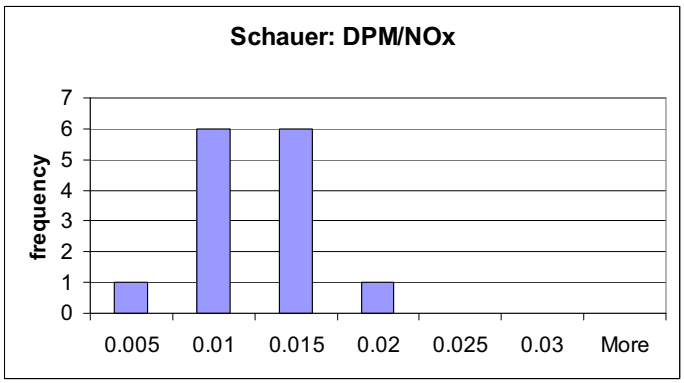


Figure 5

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. 5. 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).

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 $\mu g/m^3$ (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%). 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 $\mu g/m^3$ (based on the limited data), and the average FEC at Roseville is approximately 0.65 $\mu g/m^3$, 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 $\mu g/m^3$ and Sacramento as 0.8 $\mu g/m^3$, and the statewide average as 1.0 $\mu g/m^3$. These estimates, however, differ by over a factor of 2 from the recent MATES III organic marker CMB estimate of >3 $\mu g/m^3$ in 2004-2005 (in Wilmington). Therefore, while these two independent estimates, yielding approximately 1 $\mu g/m^3$ ambient diesel PM (in the South Coast air basin), provide further support for the lower end of DPM/ NO_x ratio, considerable uncertainty remains in their interpretation (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 into and 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

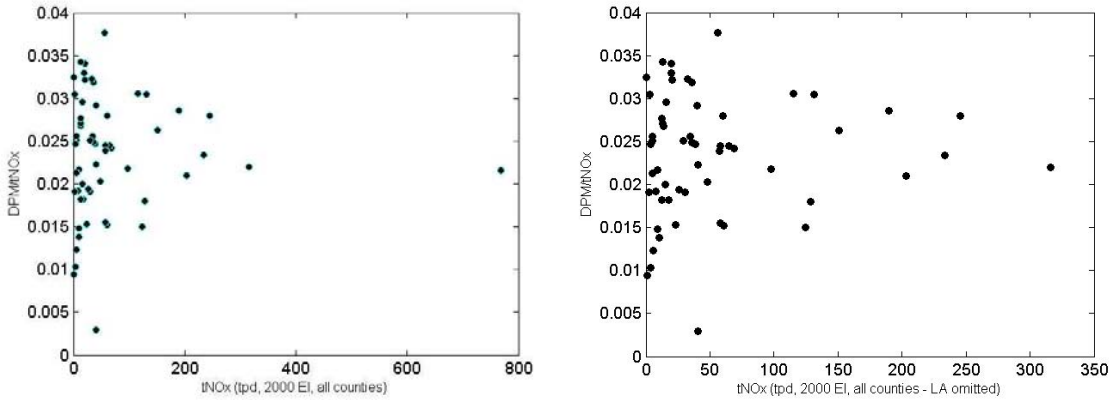


Figure 6

toward higher values, the mean residence time of an air parcel in a coastal air basin is often a few hours, which would dominate the reaction and deposition rates and effectively make the rate of removal for NO_x and PM the same. In the following, we assume this dominance of air parcel residence time on removal rates, and take the removal rates for NO_x and DPM as equal.

To compare the above source apportionment estimates of DPM/ NO_x with emission inventory estimates, we utilize ARB emission inventory estimates for the year 2000 (the SA studies were conducted in 1995 SC, 2000 SJV, and 2001 Gasoline/Diesel PM (GDPM) Split SCAB). The emission inventory estimates incorporate spatial and temporal averaging over large scales and therefore may be used to estimate average

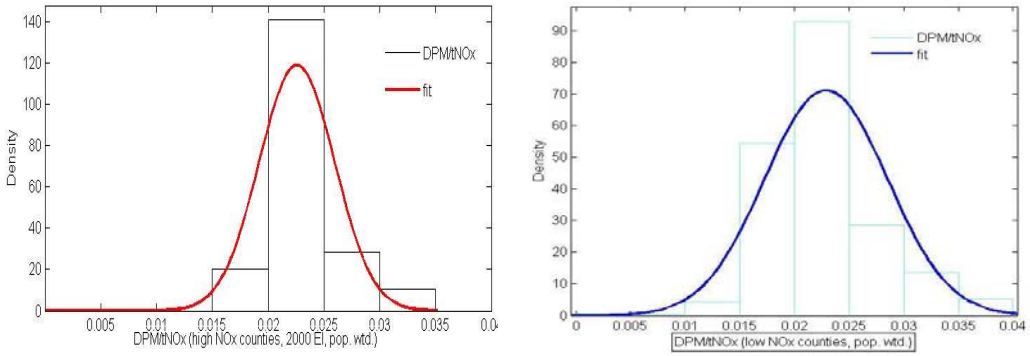


Figure 7

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)$_e$ against $NO_{x\,e}$ for all counties in California is shown in Fig. 6. 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. 6. These scatter plots show that the county-wide ratios DPM/NO_x are clustered about an average and that the dispersion depends on the average annual NO_x emission rate. The second plot in Fig. 6 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 Results section of

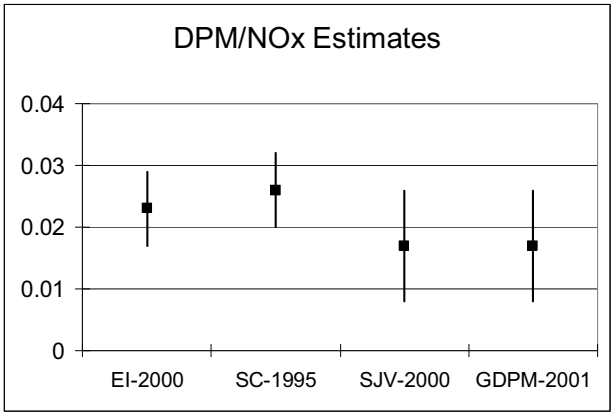


Figure 8

this document.) High-NO_x counties are highly urban and have similar composition of

diesel to non-diesel emission sources. In California, NO_x inventories are dominated by on-road and off-road mobile sources; overall, diesel engine emissions contribute approximately half of the on-road NO_x emissions and greater than 90% of the off-road mobile emissions. 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 7 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. Hence, population weighted distributions for α in high and low-NO_x counties may be described by normal distributions with same mean value and a dispersion that depends on whether the county is highly urban or not. To develop a single California-wide approximation, we take the (larger) dispersion of the ratio for DPM/ NO_x in low- NO_x counties as measure of the variability that is encountered locally within air basins.

The above estimates of the ratio DPM/ NO_x from the 2000 EI population-weighted and SA studies compare well, given the relatively large uncertainty: 2000 EI: county average 0.023 (0.006); and SA: 1995 SC 0.026 (0.006), 2000 SJV 0.017 (0.009), and 2001 Gasoline/Diesel PM (GDPM) Split SCAB 0.017 (.009); see Fig. 8. This agreement between 2000 EI and SA estimates for α motivates adoption of a single scaling factor for the whole state of California for years close to 2000: the average from the 2000 EI estimates: α = 0.023 (0.006).

Conclusions

Based on the agreement between SA and EI estimates of the scaling factor α - the ratio of DPM to total NO_x - for years close to 2000, we propose use of a single value for α in estimating the population-weighted annual average ambient DPM concentration in 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 > .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

DPM concentration estimates (µg/m^3) for year 2000			
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

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.

A comparison of DPM concentration estimates for the year 2000 using the proposed NO_x-scaling method 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% of the population of California and contribute greater than 96% 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% (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% 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.

A final application of the NO_x-scaling approach is to estimate α for a more recent year - 2005 (to be used as a baseline). An analysis of the ARB 2005 emissions inventory, similar to that carried out for 2000, yields the value $\alpha = 0.025$ (0.006). As expected, the population weighted average is close to that for 2000 - the average is 9% higher (while the standard deviation is the same). In contrast to the year 2000, however, the mean values of α in high and low-NO_x counties are slightly different: .026 (high NO_x) and .024 (low NO_x). This small NO_x dependence of α indicates that in future years a single value of α may not be suitable to describe all counties and that a further refinement of the approach, based on NO_x emission estimates, may be necessary. In any case, it is

expected that α will slowly vary with year because of changes in technology and turnover of emission sources. Although time intervals for such updates in estimates of α can not be prescribed, revisions are necessary when the difference in α estimates approaches the uncertainty (dispersion).

A rough comparison of this 2005 EI estimate may be made with measurements from recently completed field work – the Harbor Community Monitoring Study (HCMS).[11] This program was conducted in 2007 to characterize the spatial variations in concentrations of toxic air contaminants (TACs) and their co-pollutants within the communities of Wilmington, West Long Beach, and San Pedro in California's South Coast Air Basin. These communities were chosen because of the close proximity of residents to many emission sources, which include the Ports of Los Angeles and Long Beach, petroleum refineries, intermodal rail facilities and the greatest concentration of diesel traffic in the Los Angeles metropolitan area. Three types of air pollution sampling were carried out: saturation monitoring network operated by the Desert Research Institute, mobile sampling by the University of California, Los Angeles and California Air Resources Board, and a network of particle counters operated by the University of Southern California. In the interpretation of this data, several caveats should be noted. First, the high density of emission sources in the HCMS area may produce ambient NO_x and DPM concentrations that are different from those in the greater region, and hence, less well described by EI estimates. Second, source apportionment was not carried out for this study; instead, the tracer used for diesel-engine PM emissions is elementary carbon (EC) from PM2.5, which is predominantly but not exclusively from diesel emissions. Under these limitations, DPM was estimated from EC concentrations as the total carbon from diesel engine emissions (DTC): EC=.65 DTC. This conversion is adopted both for simplicity and because of the general agreement between TC and mass for diesel emissions.[8, 9] A full consideration of various methods to estimate DPM using EC, and possibly other co-pollutants, was not attempted; such an investigation would yield a confidence interval for DPM estimates, and provide a better case for comparison. The scaled-EC DPM estimates are plotted against ambient NO_x concentrations (20 sampling sites; 4

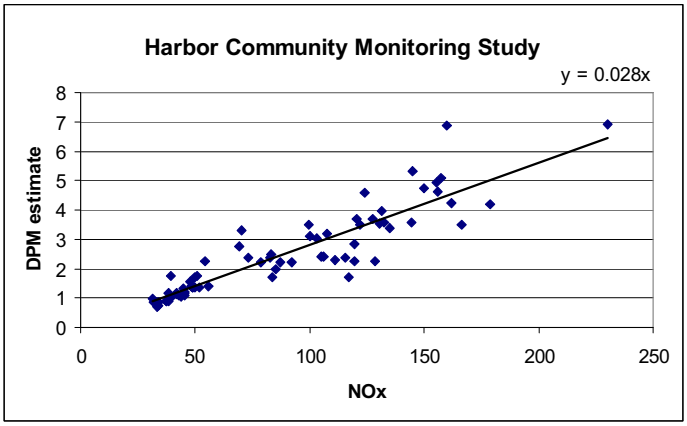


Figure 9

seasons) in Figure 9. Based on statistics for individual DPM/ NO_x ratios, the mean is 0.028 (0.006). In light of the caveats and simple EC-scaling, this result agrees very well with 2005 EI estimate of 0.025 (.006) and provides support for use of a NOx-scaling methodology (with associated uncertainty interval) to estimate DPM in California.

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 on county-wide scales, proportional removal rates for NO_x and DPM (including air basin outflow), proportionally time-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 for support. 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 mobile sources). As such, the uncertainty describes the confidence in α to accurately describe local NO_x emission sources. For areas outside California, in which the NO_x emission inventory has a significant contribution from non-mobile sources (e.g. power plants or refineries), the value of α is likely to be different from that for California and vary with source proximity. 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

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 (extraction data 7/2008).

11 California Air Resources Board – Harbor Communities Monitoring Study website (http://www.arb.ca.gov/research/mobile/hcm/hcm.htm).

Appendix 4 (Peer Review Process and Results)

A. Peer Review Process

Following the Cal/EPA External Scientific Peer Review Guidelines, CARB staff submitted a formal request to the Cal/EPA Project Director for the review of an earlier draft of the report. In it, staff clearly listed the scientific issues relevant to the proposed methodologies in the staff report and stated the required expertise in the reviewers for a successful evaluation of the proposed methodology. The scientific issues included:

1. Development of a credible range based on expert opinions from a panel of experts selected by U.S. EPA. The individual confidence intervals provided by the experts from EPA's elicitation are examined and a range of values that describe the PM2.5-mortality relationship are developed. This range consists of a mean central estimate, reflecting the best single-point estimate of the relationship between PM2.5 exposures and premature death, accompanied by low and high values that reasonably capture the uncertainty around the current state of knowledge on the relationship between PM2.5 exposures and premature death.

2. Sensitivity analysis. The robustness of estimates based on the newly developed credible range is demonstrated by comparing the results against alternative methods. One alternative is based on empirically combining the experts' distributions. Several alternative methods are considered for combining the distributions, which may include: simple averaging, inverse-variance weighting, and/or fixed random effects pooling using Monte Carlo methods. The core estimates are also compared against empirical results taken directly from the published studies.

3. Estimation of premature death associated with exposures to PM levels that do not meet health-based ambient air quality standards. A procedure similar to peer-reviewed methods used by U.S. EPA is used to estimate the overall impacts due to air pollution exposures based on new relationships that relate changes in PM2.5 to premature death. Then, in order to estimate premature deaths associated with public exposures to emissions resulting from particular sources**,** the overall impacts to particular sources (for example, locomotives or trucks) is apportioned by using the ratio of emissions associated with the source over total emissions.

Also, staff submitted a listing of individuals who may have a conflict of interest, including the scientific advisors and the experts in the U.S. EPA expert elicitation. Reviewer candidates were independently identified by the University of California at Berkeley, Institute of the Environment, in collaboration with UC colleagues. Each candidate was required to complete a Conflict of Interest Disclosure form, which was reviewed by the Cal/EPA Project Director for the independent peer review. Candidates were accepted as reviewers only if the disclosure information showed they had no conflict of interest related to the report.

The selected 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 exposure to ambient PM. The peer reviewers provided staff with written comments on the August 23, 2007 version of the report. Staff then addressed and incorporated the results of this peer review into a draft report for public release on May 22, 2008. In this report, the final methodology has taken into account all public comments received by July 11, 2008.

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

B. Peer Review Results

Results of the peer review on the general methodology are presented in section C. In addition, to clarify the application of the methodology to estimate premature deaths associated with public exposures to emissions resulting from particular sources, the peer reviewers considered the two scenarios of applying it to small areas and populations. The results of this latter review are presented in section D.

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 included in section E of this appendix.

C. Comments on General Methodology Described in the Draft Report

In this section, a summary of comments on the general methodology 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 $\mu g/m^3$ 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 $\mu g/m^3$ is not defensible. Should use 2.5 $\mu g/m^3$.
Roll-back	Reasonable. Clarify the use of background 2.5 $\mu g/m^3$.	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.

C.1 Jeffery Brook

**Scientific Review of the Air Resources Board (CARB)
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 PM2.5 health effects and PM2.5 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 PM2.5 concentrations. Above all, this requires scientific judgment, with frank discussion of the assumptions made and the limitations of the method. Overall, the CARB 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
CARB'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 CARB 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 PM2.5 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 PM2.5 exposure (as represented by an annual or multi-year average ambient concentration). The CARB 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 CARB 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 CARB 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 CARB'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 staff 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 CARB 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 CARB'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, CARB 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 CARB. Given the available information, the method developed by the CARB staff is scientifically acceptable in that multiple approaches were considered and evaluated against the recommended values. However, it is noted that CARB's range is narrower than any of these approaches. For the upper end, this implies that CARB 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 CARB 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 CARB 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

PM2.5 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 CARB). 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 PM2.5 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 CARB'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 PM2.5

The approach CARB 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 PM2.5 data are used appropriately to estimate the population exposure. Although the interpolation method used to assign monitoring site PM2.5 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. CARB 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 PM2.5, 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 PM2.5 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 CARB 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 μg/m3 for the cut-off would make much difference. However, using different values between 7 and 12 μg/m3 could affect such results.

In the second part of this section of the draft report, where CARB 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 PM2.5 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, CARB 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 CARB has set BG=2.5 μg/m3. 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 μg/m3 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 PM2.5 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 PM2.5 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.

C.2 Mark D. Eisner

Critique of "Methodology for estimating the premature deaths associated with long-term exposures to fine airborne particulate matter in California." CARB, 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 PM2.5 was discussed on page 24. One issue to consider is potential interactions between SO_X, ozone, and PM2.5. Because ozone and SO_X levels vary geographically, would the health effects of PM differ in California vs. other areas with different ozone and SO_X 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 PM2.5 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 PM2.5 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 PM2.5 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 PM2.5 cut-off value. The analysis uses a cut-off PM2.5 value of 7ug/m^3. Yet it is stated that 11/12 experts agreed that health effects may be observed at all levels of PM2.5. 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 2.5 ug/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 PM2.5 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.

C.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 PM2.5 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 PM2.5 exposures.

Question 2: Does the methodology provide a reasonable basis for the assessment of the threshold for the effect of PM2.5 *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 conceded that definitive studies needed to ascribe a threshold would be difficult or impossible.

In their considerations for the present methodology report, CARB staff 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).

C.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.

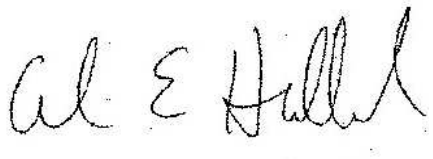
Sensitivity Analyses
I am not sure how to interpret taking the upper confidence bound from one study and lower one from other. I think a more defensible method for calculating the uncertainty bounds on the effect estimate would be a more formal method, such as those presented in the sensitivity analyses. 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

C.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.

C.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 staff 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 staff 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.

D. Comments on Application to Specific Emission Sources

In this section, a summary of comments on the application of the methodology to specific emission sources from the peer reviewers is presented, followed by individual comments from the six experts.

Summary of Peer Reviewer Comments

On CARB'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.

Issue		J. Brook	M Eisner	R Flagan	A Hubbard	J Kaufman	J Schwartz
	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-bound estimate.	No comment	The method would be applicable and needed in certain regions, esp for ammonium nitrate.	Secondary more uniform and more certain. Wood smoke and traffic likely underestimate mortality because efficient exposure.
	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	Secondary PM would be more difficult due to chemistry.	Not appropriate to use linear rollback for secondary 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.	

D.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. Methodology based on modeled concentrations

This 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 CARB'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.

D.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 CARB'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:

CARB 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

D.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).

D.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. 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.

9. Use CARB'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:

CARB estimated that the diesel PM concentration in county Y is 2 μg/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

D.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. 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?
    - 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 CARB'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.

Hypothetical Example:

CARB 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?

[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.

- 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

D.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?*

E. 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.

E.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 CARB 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 CARB 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).

Edited copy of document with Brook's comments in bold and italics:

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 (county wide). 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:

$$S = \alpha C, \qquad (1)$$

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 both emission inventory (EI) and source apportionment (SA) estimates of the scale factor α agree and that the distribution of values (over counties or measurement sites) is well approximated by a Gaussian distribution with mean 0.023 for the year 2000. This single value may be used to infer DPM concentrations from measurements of ambient NO_X concentrations; the uncertainty in estimates is based on the distribution of ratios for high- and low- NO_X emission counties.

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 (above-background) 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. The previous method[3] used by the Air Resources Board in 1998 to estimate ambient DPM10 concentrations was based on measurements of ambient PM10. In this method, one of two scaling factors, rural or urban, which were determined from chemical mass balance source apportionment studies (CMB), was used to scale PM10 concentrations to obtain DPM10 concentrations. Each DPM10 concentration was then multiplied by another factor that described how diesel emissions in the particular county

compared with the base value (rural or urban) - the ratio of the percentage of DPM10 emissions in a particular county to the base rural or urban percentage of diesel PM10 emissions. Application of this method, therefore, depends on several factors, the most important of which are: measurements of ambient PM10 concentrations, previous source apportionment work in specific air basins, and emission inventory estimates. These factors are also the primary weaknesses of the method. In particular, PM10 contains predominantly crustal material and the fraction associated with diesel PM is 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 source emissions as more recent models. We believe the proposed use of ambient NO_X concentrations is a more direct method to estimate DPM than the PM10 method, because of the 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 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 DPM to NO_X. First, we compare estimates of DPM/ NO_X ratios from two source apportionment (SA) studies with estimates based on 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. We take α to be the population-weighted EI average value of DPM/ NO_X. Second, we use the distribution of county EI values of DPM/ NO_X to estimate the uncertainty interval for α.

Two source apportionment studies, each lasting approximately one year and utilizing organic chemical speciation for chemical mass balance (CMB) apportionment of PM, have been carried out in California since 1990. The use of organic chemical markers for vehicle emissions is considered to be essential for the accurate separation of gasoline from diesel-fueled engine emissions. The first study 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 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.

The results from these two groups, J. Schauer et al. and E. Fujita et al., however, may not be directly comparable, especially in the context of the separation of diesel from

gasoline engine emissions. Recent work in which both of these groups carried out side-by-side source apportionment work in the South Coast Air Basin – the Diesel-Gasoline Particulate Split Study (2000) [7] – raised several important, but still unresolved, questions. 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. At best, similar studies by each group may be compared, differences noted, and a conversion factor(s) estimated. In the comparison of J. Schauer and DRI's methods, two recent studies in the San Joaquin Valley may be utilized: the winter intensive portion of CRPAQS (2000) by DRI and the winter study by J. Schauer as part of the Integrated Monitoring Study (1995) [8]. A section in the Final Report by DRI for CRPAQS on the uncertainty in source apportionment estimates contains such a comparison. Expressed as a fraction of total PM, the diesel contributions in the two studies are: 9.6 ± 1.1 from Schauer and 5.4 ± 1.7 from DRI, expressed as a percent of the total PM apportioned. Considering the differences in years (1995 vs. 2000), apportionment periods (6 vs. 15 days) and meteorology, the two results are relatively similar. The factor of two between the estimates likely reflects differences in ambient conditions/years in conjunction with those caused by respective methodologies; separation of these effects is not possible. Prior studies by these groups, however, show a general trend in the assignment of diesel PM, with J. Schauer's DPM estimates typically higher than those from DRI. We adopt a pragmatic viewpoint based on these prior studies that assigns all of the difference in the above SJV results to methodological differences and assume that the ratio is independent of location. Hence, we expect the ratio of 9.6/5.4 ≈ 1.8 to be a general conversion factor (DRI → Schauer) for these groups' estimates of DPM. In the remainder of this document, we make the explicit decision to use the values from Schauer et al. and translate DRI estimates to comparable results with the above conversion factor. This choice is based on compatibility with previous work in the Children's Health Study and any subsequent health studies that utilize J. Schauer's apportionment estimates. In addition, a tremendous amount of source characterization and apportionment work in California has been pioneered by G. Cass' group (including J. Schauer) at Caltech; and the results of these efforts will remain a fundamental source of information and guidance to future researchers.

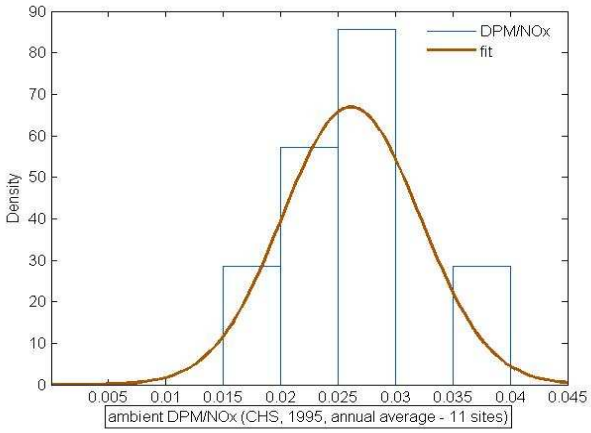


Figure 1

Relatively weak argument. DRI, clearly does quality work and they have probably done more of the recent work and certainly have more experience with CMB than

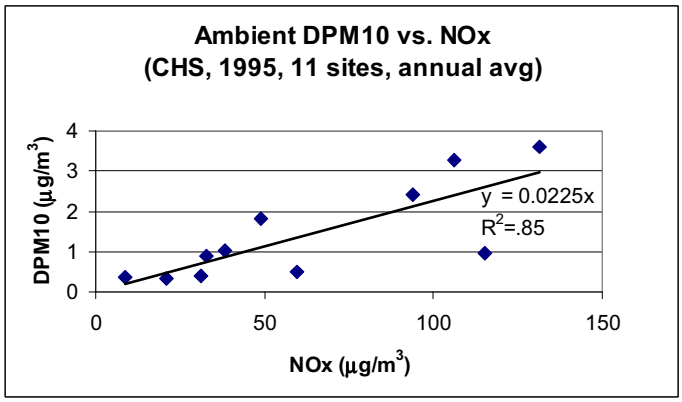
Schauer.

Results

Figure 1 shows site-to-site variation of source apportionment estimates of the ratio (annual average DPM10)/(annual average NO_X)

Where are these sites? Were the sites with the organic marker measurements and the NOx measurements the same of if not, how different were their locations?

from the CHS (1995). A straight average over all 11 sites, independent of population, gives the mean value 0.024 (0.011), where the value in parentheses denotes the standard deviation. An alternative estimate based on regression of the SA 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 is physically meaningful, as one expects that diesel emissions tend to zero with total NO_X emissions.



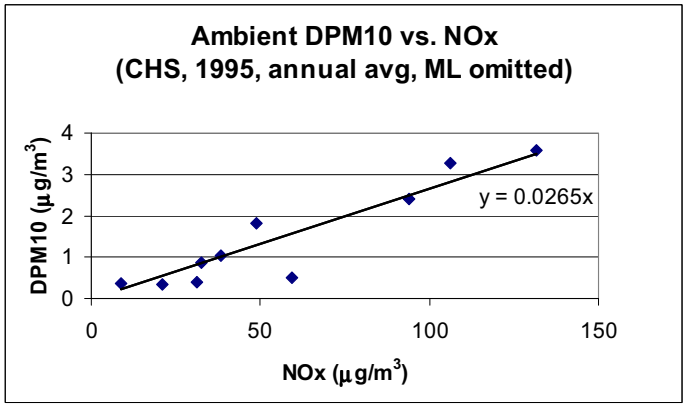


Figure 2

Not really. DPM is more likely inert while NOx reacts to NOz so with age NOx drops faster than DPM and it would be physically possible to have no NOx (especially motor vehicle NOx) and still suspended DPM.

Removal of an outlying value (for Mira Loma) gives a slope of 0.026 (0.006), which is also shown in Fig. 2.

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

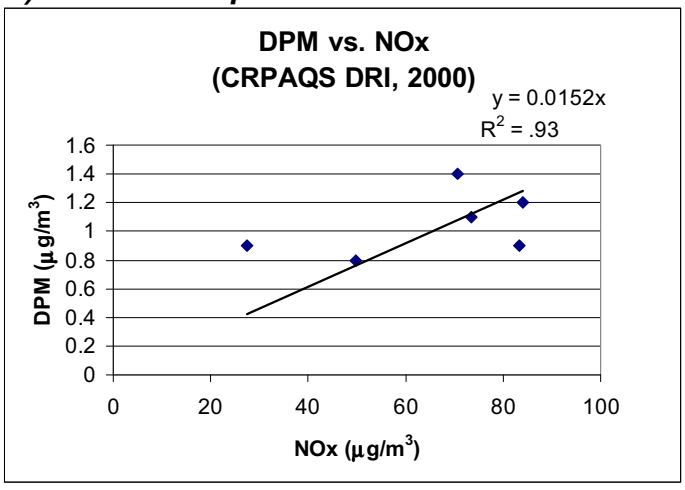


Figure 3

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, which may arise from downwind sites with low population. We believe the regression coefficient and associated standard deviation are better estimates than DPM/ NO_X statistics for DPM exposure-related work, because the most influential data in such regressions are associated with high NO_X and/or DPM, which are indicators of population density. We take the standard deviation of the slope as the measure of uncertainty in α for SA studies.

The other SA estimate for α is for the San Joaquin Valley (CRPAQS, DRI, 2000). A straight average of SA values for ambient DPM/ NO_X, for 6 sites in SJV, gives 0.017

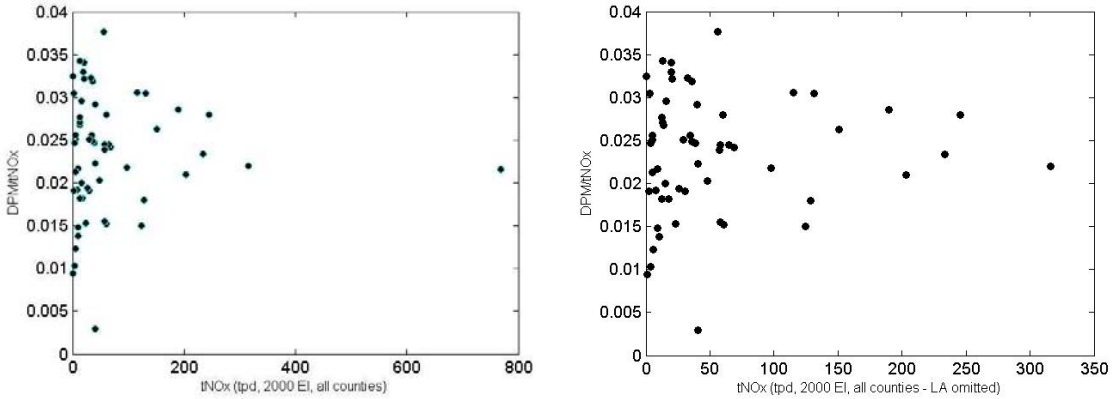


Figure 4

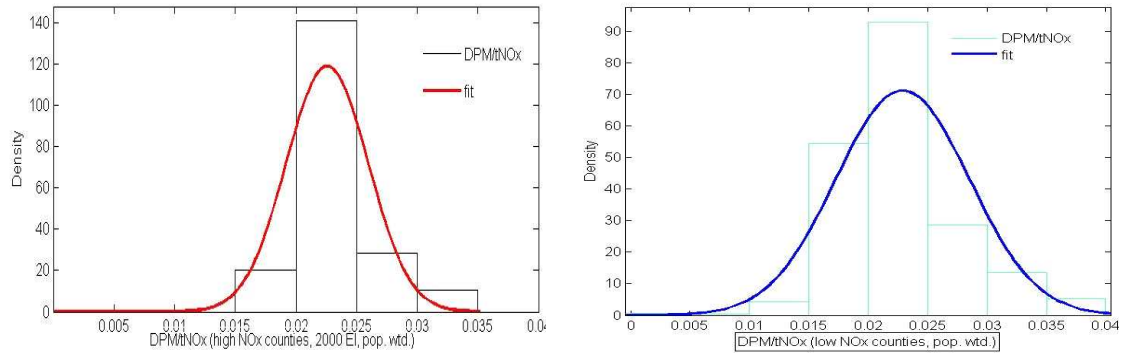


Figure 5

(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 (0.004) from the regression as an indicator of the variability in (population weighted) ambient ratios. Using the conversion factor 1.8 (Schauer/DRI), we obtain a corresponding value of 0.027 (0.008) in the context of J. Schauer's apportionment results.

A direct comparison of the above SA estimates with the emission inventory can not be made: emission inventory estimates are for whole counties while SA estimates are specific to monitoring sites and implicitly take into account meteorology, chemistry and

deposition. ***But there should be some relationship and some idea as to which ratio would be larger. I feel the ambient one should be larger.***
Assuming that the previously determined regression slopes give estimates of population weighted DPM/ NO_X ratios, we develop a comparison with population weighted EI estimates of DPM/ NO_X ratios. ***Is this traffic NOx or all NOx? I assume this is all NOx emissions.***
For this purpose, the EI estimates for DPM and NO_X emission rates from individual counties are utilized[9]. These estimates may be visualized as average tons per day of pollutants emitted into a well mixed box covering each county, with removal rates of DPM and NO_X proportionately the same. Such estimates automatically incorporate spatial and temporal averaging over large scales and therefore may be used to estimate average DPM/ NO_X ratios directly. A plot of DPM/ NO_X against NO_X for all counties in California is shown in Fig. 4. 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. 4. These scatter plots show that the ratios DPM/ NO_X are clustered about an average and that the dispersion depends on the average annual NO_X emission rate. The second plot in Fig. 4 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, and compare with the previous SA regression values, each county value is weighted by its population; weighted histograms are approximated by normal distributions. Figure 5 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. That is, 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.

The above estimates of the ratio DPM/ NO_X from the EI-population weighted and SA studies compare very well: EI 0.023 (0.003 or 0.006) and SA SCAB 0.026 (0.006) and SJV 0.027 (0.008). This agreement between EI and SA estimates for α, coupled with the uncertainty ranges, 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).

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 from NO_X

concentrations. These DPM estimates depend upon the network of ambient NO_X measurements from the CARB 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 previously determined county-specific standard deviation, i.e. high- or low-NO_X county value. The following twelve counties are considered high-NO_X (annual average NO_X > .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 areas without monitoring sites, the smoothing method may be modified to incorporate a minimum concentration (or such areas may be removed, if the population is sufficiently small). ***Explain a little more. Is implying that you expect that such areas have low concentrations?***

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 are used to project concentrations for future years.

Table 1

DPM concentration estimates (μg/m^3) by air basin			
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

A comparison of DPM concentration estimates from the proposed NO_X-scaling method (proposed) with the previous (projected) PM10-scaling method[3] (previous) is given in Table 1. The overall agreement between DPM concentration estimates is good, and for the six highest population air basins, it is very good. More specifically, the six highest population air basins contain over 90% of the population of California and contribute greater than 96% 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 20% (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 less populated 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% 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 CARB 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.

References

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 Schauer J. J. and Cass, 2000 G. R*. Source Apportionment of Wintertime Gas-Phase and Particle-Phase Air Pollutants Using Organic Compounds as Tracers*. Environmental Science and Technology. 34, 1821-1832.

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

E.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).

Appendix 5 (Public Comments and Staff Responses)

In this appendix, we summarize the key comments received from the public on the May 22, 2008 draft report, and our responses to them. Similar comments are grouped together according to the following codes:

AAM = Alliance of Automobile Manufacturers- John Heuss (Air Improvement Resources), Sacramento, CA
AIE-JS = Joel Schwartz, American Enterprise Institute, Sacramento, CA
BNSF = BNSF Railway, Topeka, KS
DTF-AS = Diesel Technology Forum – Allen Schaffer, Frederick, MD
DWN = D. Warner North, NorthWorks, Inc., Belmont CA
EMA = Engine Manufacturers Association – Joseph L. Suchecki, Chicago IL
EMA-M = Suresh H. Moolgavkar, Exponent, Inc. and consultant to EMA
EPA = U.S. EPA, Lisa Conner – Research Triangle Park, NC
FWL = Frederick W. Lipfert, Environmental Consultant, Northport, NY
HEI = Health Effects Institute, Boston, MA
IEc = Industrial economics incorporated, Cambridge, MA
JDD = John Dale Dunn, Carl R. Darnall Army Medical Center, Ft. Hood, TX
UCLA-JE = James Enstrom, UCLA, Los Angeles, CA
UCS/EDF = Union of Concerned Scientist/Environmental Defense Fund

All comments are presented in full text in a separate document called "Public Comments on Methodology for Estimating Premature Deaths Associated with Long-term Exposure to Fine Airborne Particulate Matter in California, Supplement to Staff Report".

1. **Choice of studies for draft report** - Draft report emphasized positive studies and omits consideration of negative chronic mortality studies (i.e. Veteran's study and Enstrom (2005)). In addition, many of the studies chosen were not California-centric. (AAM,AEI-JS, FWL, JDD, EMA, EMA-M)

Several commenters suggested that CARB staff examine the range of short-term studies relating PM2.5 and mortality. The focus of this report is the association between long-term PM2.5 and premature mortality. Short-term studies are used only as supporting evidence for the relationship with chronic exposure.

One commenter suggested that the Dublin coal-ban intervention study did not apply to California because the PM concentrations in Dublin are higher than in California. While CARB staff acknowledges that the Dublin study recorded higher levels of PM than in California, we think that the study lends support to the association of long-term exposure to PM2.5 and premature mortality.

Some commenters suggested that CARB put greater emphasis on the Enstrom (2006) study. CARB staff convened a teleconference with Dr. Enstrom and several prominent epidemiologists to discuss his findings. We amended that portion of the report to reflect

the discussion, which focused on two main issues: the time of follow-up since initial enrollment of the cohort, and the age of the cohort.

The first issue is the 40 year follow-up period. At first glance, this long follow-up is an attractive idea. However, the Cox proportional hazards model is influenced by long-term trends that are not likely to remain proportional to the hazard for periods of that duration, for example, changes in health care (Janes et al. 2007). While it is unlikely that changes in health care, land use, demographics and other risk factors vary on the scale of a few years, they will change over 40 years, and this is not accounted for in Dr. Enstrom's study. The original ACS and Six Cities studies were less than ten years in duration, reducing the likelihood that this issue applies to them. However, as follow-up in these populations continues, this will increasingly become an issue, unless updates to model adjustments for these factors are made.

The second issue is concerned with the age of the cohort. It is likely that at some point across a 40-year period the risk of dying in any given year dwarfs any additional risk added by PM2.5, making additional risk related to PM2.5 undetectable. As the subjects move into the older age categories, it will become increasingly difficult to distinguish additional risk from PM2.5 from that related to age. Such is suggested by Zeger et al. (2008) as well. In fact, the Enstrom paper demonstrates this, in that the relative risk for a PM2.5 effect on death decreases through the various measurement periods reported in the paper. It should be noted that Enstrom's relative risk for the 1973 to 1983 time period is similar to that reported by Pope et al. (1995) using the same exposure data, and when the subjects in the two groups were of similar ages.

One commenter suggested that CARB staff should focus on exposure to source-specific PM constituents rather than PM mass. Proximity to traffic, in particular, was suggested as the true cause of health effects. Research on the health effects associated with traffic-related constituents is an active area of investigation. However, CARB staff think that the available evidence suggests that PM2.5 mass is the best pollutant to explain the health effects associated with ambient particulate matter. The toxicity of "generic" ambient PM2.5 in any time or place is the combination of the toxic potential of all constituents of that mass of PM2.5, justifying the approach of regulating PM2.5 as a "generic" mass. Most PM2.5 epidemiology studies published to date have relied on monitoring data for "generic" PM, and the majority of them have reported a statistically significant association with mortality. Hence, it is justifiable to use PM mass as a pollutant for health impacts quantification.

2. **Epidemiologic associations (temporal) -**The overall pattern of the epidemiological associations in the chronic and acute studies is not consistent with a generic ambient PM2.5 mortality signal. In the draft report, it is assumed that the association between PM2.5 and mortality is constant over time and that the temporal trends are the same nationally and locally. This assumption does not always hold (Janes et al, 2007). (AAM, BNSF, EMA-M)

This comment has several parts. The section dealing with acute exposure to PM2.5 is

not directly relevant to this staff report, in that the report addresses a methodology for estimating the influence of long-term PM2.5 exposure on mortality, and studies on acute exposure are used only as supporting evidence for long-term exposure.

It is unclear why the commenters emphasize “generic” PM2.5. While it is true that the chemical composition of PM2.5 varies temporally and spatially, toxicological research suggests that all particles have some degree of toxic potential, although some seem to be more toxic than others. The toxicity of “generic” ambient PM2.5 in any time or place is the combination of the toxic potential of all constituents of that mass of PM2.5, justifying the approach of regulating PM2.5 as a “generic” mass. Most PM2.5 epidemiology studies published to date have relied on monitoring data for “generic” PM, and the majority of them have reported a statistically significant association with mortality.

The Cox proportional hazard model assumes that the association between PM2.5 and mortality remains constant over time. This is likely a reasonable assumption over periods of several years, although it would not be a reasonable assumption to apply over long periods of time. Schwartz et al. (2008) found that effects of changes in exposure on mortality are seen within one or two years. Thus, it is appropriate to apply results from studies that use the Cox proportional hazard model. Violation of this model assumption will have an increasingly important effect the longer a cohort is followed, which would have the effect of increasing uncertainty. One of strengths of the Cox model is that it exploits the spatial differences in PM2.5 exposure between areas, allowing comparison of mortality rates in areas with different PM2.5 concentrations.

The Janes et al. (2007) paper is a report of methods development that is subsequently demonstrated. The study is actually a time-series model, although it is based on monthly PM2.5 mean concentration, rather than the usual daily mean. This approach limits the ability to exploit daily PM2.5 variability, and by including county-specific indicators, the ability to exploit spatial differences in PM2.5 concentrations. These factors reduce both long-term and daily PM2.5 variability, leaving limited exposure variability between communities or areas to exploit in examining the relationship between PM2.5 and mortality. The stated goal of the paper was to examine the influence of sources of long-term variability that could confound the relationship between PM2.5 and mortality. Given the short time span of the study (1-3 years, depending on the analysis), it is unlikely that the potential sources of long-term variability specified as being concerns, such as changes in mortality rates, industrialization, health care, land use, or large scale weather events, would change enough to have much confounding influence on the outcome of the study. It is also interesting to note that when risks were estimated based on the previous year’s exposure there was considerable variability among the three years of the study. The effect estimate was also strongly affected by the number of degrees of freedom included in the model.

3. **Diesel specificity** - CARB methodology assumes that all components of PM2.5 are equally toxic. This is not consistent with the large body of toxicological data on either individual PM2.5 components or ambient PM2.5 mixtures. In addition,

there are no observational or controlled exposure studies which demonstrate that diesel PM2.5 is equally toxic to generic PM2.5. Based on these facts, the new methodology is not relevant for control of diesel emissions. It was suggested that the current draft report be broken into two separate reports: one for ambient PM2.5-associated health effects, and another for estimation of ambient diesel PM. (AAM, EMA)

CARB staff agrees that research needs to be conducted to explore the relative toxicity of diesel and other sources of PM. In the absence of additional information, staff can only assume that all components of PM are equally toxic. As discussed in the report, the extensive animal toxicology literature on the health impacts of diesel PM exposure leads to the conclusion that diesel PM is at least as toxic as the general ambient PM mixture.

There was a suggestion to make two separate reports, one for DPM and another for ambient PM. Since diesel PM contributes to ambient PM, and in light of the lack of results to suggest that diesel PM is less toxic than ambient PM (CARB 1998), CARB staff think it is justified to include estimates of DPM and ambient PM2.5 in the same report.

One commenter cited a study of atherosclerotic prone mice exposed to concentrated ambient particles from Tuxedo and Manhattan, NY, with no effect for the high DPM exposure case. The study cited has not appeared in the peer-reviewed literature. As new peer-reviewed literature emerges, CARB staff will revisit PM-associated mortality.

A commenter wrote that diesel PM could not be as toxic as ambient PM based on the EPA reference concentration (RfC) of 5 $\mu g/m^3$ for diesel exhaust particles. The EPA RfC is based on animal studies that demonstrate evidence of chronic respiratory effects. The animal studies show consistent findings of inflammatory, histopathological (including fibrosis), and functional changes in the pulmonary and tracheobronchial regions of laboratory animals, including the rat, mouse, hamster, guinea pig, monkey, and cat. These animal studies show inflammatory response which is consistent with the prevailing hypothesis for how PM2.5 causes cardiovascular disease. Also, the scientific literature for the EPA RfC is generally current through January 2000, although a few later publications were included. Our report includes more current epidemiologic literature on the health effects of ambient PM on humans. As discussed above and in our report, diesel PM is a component of ambient PM and assuming diesel PM to be as toxic as general ambient PM, staff thinks it's justifiable to include diesel PM effects in the report.

4. **Diesel contribution to PM2.5 -** There is no scientific evidence that elimination of any specific, individual air pollution constituent, even if that were possible, would result in the prevention of the estimated number of deaths. In addition, CARB staff relies on un-validated quantitative estimates of number or deaths due to diesel particulate matter arising from goods movement or avoided by reductions in diesel particulate matter from new controls on goods movement. In fact, studies have shown the levels of diesel PM in CARB emissions inventory have been steadily declining due to new cleaner fuels and technology. This in

inconsistent with the proposals to increase risk attributable to diesel exhaust (BNSF, DTF-AS)

It is in fact possible to eliminate specific sources of PM. Intervention studies provide evidence that elimination of specific PM sources result in reduced mortality and hospitalizations (Hedley et al. 2002; Clancy et al. 2002; Pope et al. 1989, 1991, 1996). There is also evidence that reduced ambient PM is associated with reduced mortality (Laden et al., 2006). The Laden et al. (2006) study provides evidence for validating the PM2.5-related mortality.

It is possible to have an increase in the estimated premature mortality despite a declining emissions inventory. Premature mortality is a function of four factors: mortality incidence rate, population affected, PM2.5 concentration, and the concentration-response function. While the estimated DPM concentration has declined, the population of exposed people has increased over time, and in CARB's judgment, the previous estimates based on the C-R function of 6% are underestimated (compared to our current estimate of 10% per 10 $\mu g/m^3$). These factors contribute to a higher estimate of mortality compared to previous estimates.

5. **Establishment of causality (concentration-response**) - The report fails to address discrepancies regarding relationships between exposure and effects. More specifically, what effects are due to low level exposure versus peak exposure, and what biological mechanisms have been identified for the presumed effect. (AAM)

The studies reviewed in the report examined exposure to PM levels over a wide range (as low as background of 2.5 $\mu g/m^3$ and as high as 27 $\mu g/m^3$). These studies found significant relationships between exposure and the increased risk in premature death. There are mechanistic studies that provide some insight into how PM2.5 exposure might cause health effects, which are now discussed in the report. What the long-term epidemiological studies of the relationship between PM2.5 and mortality find is that people who live in areas with higher annual average PM2.5 concentrations have a greater risk of dying than people who live in areas with lower annual average PM2.5. Using the risk estimate from the staff report as an example, people who live in an area with an annual average PM2.5 concentration of 25 $\mu g/m^3$ have a 10% greater risk of dying in a given year than people who live in an area with an annual average PM2.5 concentration of 15 $\mu g/m^3$. This does not imply that people are dying from PM2.5 exposure. It means that PM2.5 exposure represents an additional risk that is added to all other risks of dying. The majority of people die from chronic diseases in the US, and the PM2.5 epidemiological results indicate that elevated PM2.5 exposure can increase the risk of dying from those chronic conditions. With a large population, the relative risk approximates a rate. Based on application of a relative risk of 10%, we estimate that chronic exposure to PM2.5 is a contributing factor in about 8% of all cause mortality in California.

6. **Establishment of causality (confounding factors) –** The draft report fails to address the effect of all known-pollution factors (weather, smoking, diet) that

might be correlated with pollution levels and health outcomes. In addition, the application of epidemiological study in one population may not be valid given possible differences in sociodemographic composition, land use, geography and meteorology between different populations. Also, the composition of PM2.5 is California is different from that in other parts of the U.S., so the toxicity may also be different. Janes et al. 2007 suggests that PM2.5-mortality associations are due to confounding factors (e.g. demographic changes, lifestyle factors), but this study was not acknowledged by CARB staff. (AAM, AIE-JS, BNSF, EMA-M)

This comment has several parts. One of the criteria for evaluation of papers was the extent to which they accounted for possible confounders and effect modifiers. All papers in Table 1 of the staff report included control for smoking, several demographic factors, and other relevant covariates at entry into the study. The Krewski et al. (2000) reanalysis of the Six Cities and ACS studies included a number of sensitivity analyses that not only replicated the original analyses, but also included evaluation of alternative models that included control for a larger number of possible confounders and effect modifiers. As shown on pages 138 and 158 of the Krewski report, once certain basic covariates were included, addition of other covariates had little effect on the resulting relative risks. Based on these extensive sensitivity analyses, we concluded that the studies cited had identified and adjusted for the important confounders.

Janes et al. (2007) investigated methods of removing long-scale temporal confounding, for example for such things as changes in health care, industrialization, or land use patterns, from long-terms exposure studies. However, their study covered only a three year period, during which it is unlikely that significant changes would occur on the national or local scale in these factors. This study is discussed in the staff report.

Our method development focused on U.S. studies. It is unclear why the commenters opine that these studies might not be applicable in California. Some have conjectured that there might be differences in the population of California compared to other parts of the country. It is true that the long-term epidemiological studies have not included large numbers of low-income or minority subjects. However, what is known about the biological mechanisms of air pollution health effects clearly shows that the same biological processes are involved in all subjects who have been studied. There is no reason to think that people of one sub-group would respond differently or through different mechanisms than another, although effect magnitude will vary among individuals. The population of California is as diverse racially, ethnically, age-wise, and socioeconomically as the rest of the US, and would be expected to respond similarly to people in other parts of the country. The Krewski et al. (2000) reanalysis shows that once age, education and smoking are accounted for, introduction of further covariates has little effect on relative risk.

Some commenters have suggested that composition of the PM2.5 in California is different from that in other parts of the US, and so the toxicity is also different. The toxicological literature indicates that particles, independent of their chemical composition, have toxicity to varying degrees. The toxicity of any given mass of PM2.5 is a composite of the toxicities of all the constituents. PM2.5 composition is not constant

on either a temporal or spatial scale at any location, making the contention that it is substantially different in one location compared to another difficult to assess.

7. **Concentration-response** – There is still uncertainty in the shape of the concentration-response function especially with respect to the presence or absence of threshold concentrations for adverse health effects. CARB should therefore include 0% in the credible range. (AAM)

This issue has been extensively addressed in the CARB/OEHHA review of the PM standard in 2002. Our review went through several public hearings as well as independent scientific peer review by the Air Quality Advisory Committee. It was also addressed in detail in the Critical Review by Pope and Dockery (Pope and Dockery, 2006). Basically, we support the scientific consensus that there is little evidence for a threshold for either the short- or long-term exposure studies. Most analyses that have specifically examined the CR functions have failed to reject the hypothesis of no threshold. In addition, in the U.S. EPA expert elicitation, most of the experts did not include zero in their lower bound estimates. Therefore, though there is uncertainty in the shape of the function, our assessment is supported by the overwhelming weight of evidence based on the recent peer-reviewed literature for a statistically significant relationship.

8. **Establishment of a PM 2.5 Mortality value –** It is not valid to calculate specific numbers of deaths at a specific site due to specific air pollution constituents even if population, location, exposure and outcome differences can be accounted for. The calculations in the draft report are based on the assumption that the relationship between individual pollutants and the health outcome is independent of all other risk factors for the same disease in the target population. Inferring that the relationship is directly casual is an oversimplification. (AAM, BNSF)

This comment has several sections. The commenters question the application of a relative risk developed for ambient PM2.5 to estimate risk associated with one constituent of the total mixture, in this case, diesel particulates (itself a complex mixture of pollutants). As noted elsewhere in these responses, the toxicity of ambient PM2.5 represents a composite of the toxicities of all components. Few data exist on the health impacts of most PM2.5 components at a whole animal or human level. However, indirect evidence from traffic studies suggests that diesel PM is at least as toxic as the general mixture of PM2.5 (Hoek et al. 2000, Tsai et al. 2000, Laden et al. 2000), and so our methodology assumes that diesel particulate is equally as toxic as the ambient PM2.5 mixture as a whole. Admittedly, this approach could underestimate or overestimate mortality associated with diesel particulate. However, the methodology is being used to compare alternative regulatory scenarios. The output from these analyses should be interpreted in this light. So long as the assumptions for the scenarios are clearly stated, the methodology can be used to compare current conditions to those that could result from adoption of a specific control regulation to obtain an estimate of the mortality risk associated with the change in PM2.5, or some component of PM2.5. The assumptions as to other components of risk, such as smoking, age, socioeconomic status, are accounted for in the models. As we have noted elsewhere in these

responses, we have not claimed PM2.5 exposure as a sole cause for premature mortality, only that PM2.5 exposure adds an increment to all other risks of dying. There are mechanistic studies that provide some insight into how PM2.5 exposure might cause health effects, and the potential mechanisms are discussed in the staff report. However, biological mechanisms don't need to be completely determined to infer that reducing exposure will lead to improving public health.

9. **Estimation of ambient PM2.5**– Because of possible future differences in source control and gasoline/diesel ratios, the DPM/NO_X may change. The methodology for estimating ambient concentrations of PM2.5 from diesel-fueled engines may also have to be adapted and should undergo a separate technical review. (AAM, BNSF, EMA)

We agree that α, the ratio of diesel engine emitted PM to total emitted NO_X, will likely change because of improvements in technology and turnover rates of emission sources. We will revise the report to address the time dependence of α by analyzing emission inventory estimates of α for two years: 2000 (for comparison with source apportionment study estimates) and 2005 (for use as a baseline). These results show a slight shift in the population weighted average ratio α from .023 (2000) to .026 (2005). Revisions such as this are necessary when the change in α estimate approaches the uncertainty (standard deviation for α).

A review of the approach used to estimate DPM based on NO_X concentrations has been conducted by several independent researchers outside of CARB: Professor Jeffery Brook of Environment Canada and University of Toronto, Professor Richard Flagan of California Institute of Technology, and Professor Phil Hopke of Clarkson University.

10. **Choice of peer reviewers (CARB)** – Despite the wide range of scientific opinion on the validity of observational epidemiology studies and air pollution epidemiology in particular, CARB Staff chose epidemiologists who are in support of the method and results in the draft report. The report therefore did not receive an independent, critical review. It has been suggested that CARB staff should commission an independent review of its analysis. (AIE-JS)

Staff worked closely with three advisors from academia and government and based the proposed methodology on the U.S. expert elicitation results. The three advisors are well respected by the scientific community. U.S. EPA contracted with Industrial Economics, Inc., due to the company's extensive experience and knowledge on conducting elicitation. In this effort, the twelve experts were chosen based on a two-phase nomination process designed to obtain a balanced set of views. The U.S. EPA expert elicitation results underwent a rigorous peer review process.

In addition, CARB's proposed methodology was subjected to an independent peer review process, with reviewers chosen by the University of California at Berkeley, Institute for the Environment. Following the Cal/EPA External Scientific Peer Review Guidelines, CARB staff submitted a formal request to the Cal/EPA Project Director for the review of an earlier draft of the report. In it, staff clearly listed the scientific issues

relevant to the proposed methodologies in the staff report and stated the required expertise in the reviewers for a successful evaluation of the proposed methodology. Also, staff submitted a listing of individuals who may have a conflict of interest, including our scientific advisors and the experts in the U.S. EPA expert elicitation. Reviewer candidates were independently identified by the University of California at Berkeley, Institute of the Environment, in collaboration with UC colleagues. Each candidate was required to complete a Conflict of Interest Disclosure form, which was reviewed by the Cal/EPA Project Director for the independent peer review. Candidates were accepted as reviewers only if the disclosure information showed they had no conflict of interest related to the report.

Six reviewers were identified by UC Berkeley and selected by the Cal/EPA Project Director to review the proposed methodology. 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.

11. **Choice of Expert Elicitation (EPA) –** CARB places a great weight on EPA's Expert Elicitation results. However of the 12 experts on the panel, six are co-authors of the studies considered, meaning they were giving their expert opinions on their own research. Also, those who find negative results are less likely to be published, so the Expert Elicitation process is biased toward researchers who have found positive results. Several members of the panel were not epidemiologists, even though they were primarily being asked to review epidemiological literature. EPA's process created subjective meta-analyses, to which CARB staff added its own additional level of subjective meta-analysis; this is not scientifically justifiable. Therefore, CARB staff should not place so much of an emphasis on their opinions. (AIE-JS, UCLA-JE, JDD, DWN, EMA, EMA-M)

The EPA panel was selected through a nominating process by Industrial Economics, Incorporated for the U.S. Environmental Protection Agency. Nominators were chosen based on a count of peer-reviewed publications. The authors with the greatest number of publications with first, second or last author were asked to provide nominations. Neither CARB nor EPA decided who would be on the panel nor who would nominate panel members.

Commenters noted that publication bias skews CARB's analysis. Publication bias may arise from researchers more likely to seek publication of a positive result, and journal editors may be more likely to accept for publication articles with positive results. However, the experts who participated in the U.S. EPA's elicitation could freely choose any study for consideration in their evaluation of the literature, and as indicated in our staff report, all studies, regardless of their results, have been carefully considered as part of the evidence in our evaluation. While CARB staff acknowledge that publication bias may occur, we will continue to rely on the best available science in peer reviewed journal articles.

12. **Pope/American Cancer Society (ACS) study** – In the draft report, CARB cites the ACS study as strong evidence that any amount of PM in the air is deadly. The study does not consider important confounding variables such as migration, or potential effects of co-pollutants. Reanalysis by HEI (Krewski et al 2000) with these confounders refutes the previous results. In addition, the HEI document illustrates (pg 197) that there is no California specific epidemiological evidence that indicates excess mortality risk due to PM2.5 among the APS CPS II cohort during 1982-1989. Both of these facts are not acknowledged by CARB staff. In the Jerrett (2005a) study, the findings that a higher relative risk was found in the Los Angeles when compared to the broader ACS analysis is intriguing, and raise the possibility that better estimation of exposure as was attempted in this study could reduce exposure measurement error and result in a "truer" estimate of risk. To test that, however, the investigators also attempted a similar analysis in the New York Metropolitan Area, which has recently completed HEI Peer Review, was presented at the HEI annual conference, and is now going "into press" at HEI. Using same ACS CPS II database and proportional hazards methodology used in Pope et al. study, ARB should calculate all cause mortality relative risk and 95% confidence interval associated with a 10 $\mu g/m^3$ increase in PM2.5. (AIE-JS, EMA-M, UCLA-JE, HEI).

Some comments are focused on Figure 21, page 197 of Krewski et al. (2000) suggest a misunderstanding of the figure. The figure is a visual overlay of the mortality and the PM2.5 surfaces as spatially modeled in one of the ACS sensitivity analyses. The figure shows that in California, the majority of the most populous regions have low to medium levels of PM2.5, and medium mortality. The exception is the Fresno area, and moving east into the Sierra Nevada Mountains. The description of the figure is on page 198, and states: "For the medium levels of pollution, intersections exist (referring to the two spatial surfaces) for high and medium mortality rates, but not for low mortality rates. Only the low fine particle category intersects with the low mortality rate category." The point of the figure was to investigate the spatial concordance between high PM2.5 and high mortality areas, not to make a statement as to specific risk in any area of the country.

We appreciate the commenter's suggestion for calculation of California-specific relative risks using the ACS CPS II cohort data. However, CARB staff does not own or have access to this data, and consequently can not perform the requested calculations. While CARB has funded projects that use the CPS II data, the agency has no role in obtaining the necessary data. In terms of studies on the relationship between long-term exposure to PM2.5 and mortality, recent research (Jerrett et al., 2005a) into spatial variability in PM2.5 concentrations across regions, for example the Los Angeles area, shows that exposure assessments based on county level monitoring, as used in Enstrom (2005) and the various Pope et al. papers (1995, 2002, 2004), do not adequately represent population exposure, and introduce a bias toward the null. Consequently, we question the utility of an analysis that relies on what is not currently viewed as the best exposure estimation methodology.

We appreciate HEI bringing to our attention the study by Dr. Michael Jerrett comparing New York City to Los Angeles. However, since the results of that study have not been published in a peer-reviewed journal, the study was not included in our analysis because we are including only peer-reviewed publications in our report.

13. **BenMAP analysis** – It is unclear to determine precisely how relative risks were calculated using BenMAP; a simple median which is not recommended due to reservations surrounding the estimate regarding causality, functional forms and distributions of uncertainty or a second approach using a Monte Carlo technique employed by EPA's contractor, Industrial Economics. (EPA-LC)

As detailed in the revised report, the medians were obtained by applying a California PM2.5 data set to the twelve expert distributions, taking into account the distributions that are conditional on a causal relationship and the distributions that vary with PM2.5 concentrations in BenMAP.

14. **Animal Studies do not support PM2.5's mortality effects –** Currently there are no animal studies which demonstrate that ambient PM increases mortality. The animal study cited by the draft document (Sun et al. JAMA 2005) used a mouse strain predisposed for heart disease. (AIE-JS)

Taken as a whole, the literature on animals and humans suggests that exposure to PM2.5 for a few hours have little adverse effect. Daily changes in PM2.5, as investigated in time-series studies, have generally been associated with a small, statistically significant, increase in risk of mortality and various morbidities. Studies of long-term (i.e. years) exposure to PM2.5 suggest a larger increase in risk of mortality. This body of literature suggests that long exposure is needed to significantly influence risk of mortality.

The commenter is correct that there are no studies in the literature in which animals have been exposed to PM2.5 for a lifetime, simulating the human condition. Consequently this literature is not helpful in supporting or refuting causality for the relationship between long-term exposure to PM2.5 and mortality. The paper by Sun et al. (2005) does provide some support for a causal relationship between exposure to PM2.5 and mortality based on the finding of accelerated progression of cardiovascular disease in an animal model prone to develop atherosclerosis. While the study did not investigate the influence of PM2.5 exposure on mortality in these animals, the finding that the animals had accelerated progression of atherosclerosis is relevant to humans, in that PM2.5 exposure has been associated with an increased risk of mortality primarily for cardiovascular diseases.

15. **Estimate range** –Given the lack of empirical data regarding a PM2.5 concentration threshold, CARB's use of a range of estimates including background levels to the lowest level of PM2.5 from the ACS study is an acceptable interim solution. As more information becomes available, the methodology should be updated. (EDF/UCS)

Subsequent to the peer review of the draft report, several new publications appeared in the peer-reviewed literature. CARB staff has carefully reviewed the recent evidence and revised the discussion of the cut-off level in the report. In light of this new information, staff feels that a cut-off level of 7 $\mu g/m^3$ would be too high and therefore recommends a cut-off level of 5 $\mu g/m^3$ We will continue to review the literature and update the methodology and assumptions as necessary.

16. **2005 Enstrom Paper –** CARB Draft Staff Report mischaracterizes Dr. Enstrom's paper (Inhalation Toxicology 17:803-816,2005). The methodology in that paper is consistent with those used in the 2002 Pope study. In addition the 2006 response to the 2006 Brunekreef criticism of the 2005 Enstrom paper was omitted. It is the largest and most detailed study ever published on PM2.5 and mortality in the California population and should be included in the draft report. CARB Staff have failed to address the important points made in four pages of public comments submitted to CARB on April 22, 2008 regarding the Goods Movement Emission Reduction Plan and the health effects of diesel emissions. (UCLA-JE, FWL, JDD, EMA, EMA-M)

In the revised report, we modified the discussion of Enstrom (2005) and Brunekreef (2006) critique, and included Dr. Enstrom's response to the Brunekreef critique.

These commenters appear to propose that the Enstrom paper should be the sole paper used to estimate PM2.5-related mortality in California, and that because it was not so selected, it was not considered in developing the revised methodology. This is incorrect. While it is true that the paper is the largest study of its kind ever conducted in California, there are several factors that led to less weight being given it than some other papers. The most significant reason for not giving greater weight to the Enstrom paper is the 40 year follow-up. At first glance, this long follow-up is an attractive idea. However, the Cox proportional hazards model is influenced by long-term trends that are not likely to remain proportional to the hazard for periods of that duration, for example, changes in health care. This is suggested by Janes et al. (2007). While it is unlikely that changes in health care, land use, demographics and other risk factors vary on the scale of a few years, they will change over 40 years, and this is not accounted for in Dr. Enstrom's study. The original ACS and Six Cities studies were less than ten years in duration, reducing the likelihood that this issue applies to them. However, as follow-up in these populations continues, this will increasingly become an issue, unless updates to model adjustments for these factors are made.

It is very likely that at some point across a 40-year period the risk of dying in any given year dwarfs any additional risk added by PM2.5, making additional risk related to PM2.5 undetectable. As the subjects move into the older age categories, it will become difficult to distinguish additional risk from PM2.5 from that related to age. Such is suggested by Zeger et al. (2008) as well. In fact, the Enstrom paper demonstrates this, in that the relative risk for a PM2.5 effect on death decreases through the various measurement periods reported in the paper. It should be noted that Enstrom's relative risk for the 1973 to 1983 time period is similar to that reported by Pope et al. (1995) using the same exposure data, and when the subjects in the two groups were of similar ages.

Another concern with the Enstrom paper is the assignment of Los Angeles as the referent city, where about half of the subjects lived (N=17,340), and where PM2.5 levels were among the highest of the counties included in the analysis. The next largest group lived in Alameda County (N=4294). Several other counties had fewer than 1,000 subjects each. The effect of this is that there would be a robust estimate for the referent county due to high statistical power related to the large population, but uncertainty in comparisons between Los Angeles and the other counties would increase as the comparison population decreased due to low statistical power in the comparisons. Thus, it would be difficult to distinguish effects due to the large uncertainty ranges in many of the comparisons. Admittedly, the clustering of about half of the cohort in one county, and the non-uniform distribution of the other subjects throughout the state make it difficult to select a referent county from this group.

Regarding the suggestion to use the Enstrom (2005) publication as the sole source for the PM2.5 concentration-response function, it should be noted that the scientific method and the development of science-based public policy require consideration of all relevant available data. Standard scientific practice is to consider the range of data reported in the literature, along with its uncertainty, and draw the most reasonable conclusion possible. When multiple studies exist, standard scientific practice does not support reliance on only one study.

Table 1 of the Staff Report lists the relative risks for several long-term PM2.5 exposure studies. The relative risks for a 10 $\mu g/m^3$ change in PM2.5 range from 0.3 (Lipfert et al., 2003) to 27 (Laden et al., 2005) for those studies conducted in the US. In light of this range, it is not scientifically reasonable to select a single paper's value, particularly one at the bottom of the distribution of reported risks, as the commenters urge. Consideration of the range and distribution of relative risks among the available studies led to the decision to select 10% (derived from the U.S. EPA expert elicitation results) as a reasonable value for subsequent analyses.

We appreciate the comments on CARB's Goods Movement Emission Reduction Plan. Since those comments were not submitted with reference to this staff report detailing the new methodology for PM mortality estimation, they will not be addressed here. They will, however, be addressed along with all other comments received on the goods movement issue as part of that regulatory activity.

17. **Median Estimate Calculations –** CARB staff should provide more detail about the process for developing the median estimates for the 5^{th}, 50^{th} and 95^{th} percentile for the PM-mortality relationship, including how CARB staff dealt with the issue of addressing the distributions that are conditional on a causal relationship and the distributions that vary with PM2.5 concentrations. (IEc)

As detailed in the revised report, the medians were obtained by applying a California PM2.5 data set to the twelve expert distributions, taking into account the distributions that are conditional on a causal relationship and the distributions that vary with PM2.5

concentrations in BenMAP.

18. **Baseline Mortality** - Baseline mortality is not defined in the report. For example, CARB staff need to clarify whether accidental deaths were included in the baseline mortality incidence data as is implied in the current discussion. (BNSF)

Baseline mortality, as used in the report, was the same as all-cause mortality, which includes accidental deaths.

19. **Premature Mortality (definition)** - Premature mortality is not defined in the report. This definition must be provided, including a definition of de minimus levels. Additionally, CARB staff should document whether premature mortality was defined consistently across the key studies, and whether the study definitions in turn were consistent with CARB's own definition. (BNSF)

Premature mortality is a death which occurs at a younger age than would be expected, as compared to baseline mortality. Air pollution is not implicated as the *cause* of death, but rather a contributing factor in someone whose health is typically already compromised, thereby accelerating the time of death. Premature mortality was used consistently throughout the studies, including the CARB report.

With respect to de minimus levels: this varies from study to study, depending on the makeup of the subject pool, the research details, etc. Therefore, it is not possible to set a fixed de minimus level.

20. **Premature Mortality (threshold determination) -** The report concludes that premature mortality may occur at concentrations below the current California standard and at or near background levels. CARB's methodology is intentionally conservative for rulemaking purposes and is particularly inappropriate for site specific analyses. In essence, even small additions to local PM2.5 levels from new/existing facilities will yield increased estimates of death and this in turn may effectively preclude industrial development and expansion. An economical impracticality in a heavily populated and industrialized state such as California. (BNSF)

The report updates the methodology for estimating premature mortality associated with long-term exposure to PM2.5 using the best available science. Policy discussions, such as industrial development and expansion are beyond the scope of the report.

21. **Policy Implications of the Draft Report –** The methods and results contained within the draft report should be used as guide for policy makers and not as a de factor regulation. It is not appropriate for a state agency to require application of and adherence to a specific method of health effects assessment outside of the regulatory process without the benefit of legislative review. Further, the ARB has not defined a *de minimis* level of premature mortality, and ARB needs to explain how the report concludes that premature mortality may occur at concentrations

below the current California standard of 12 μg/m^3 and may occur at or near background levels of $PM_{2.5}$ (2.5 μg/m^3). (BNSF)

The methods and results in this report are not required in a regulation. Hence, they are not subject to the Administrative Procedure Act (APA) and review by OAL in that they do not fall within the definition of a regulation. However, the report is a reference that we use in our regulations; thus, the methodology has been released for public review, and any reports emanating from its use will be made available for public review under the requirements of the APA.

With regard to a *de minimis* level of premature mortality, we agree that the communication of a very small risk is an important issue which should be considered when the methodology for assessments of small populations is further developed.

On the last point relating the cut-off level in relation to the standard, new publications have been reviewed that indicate that associations between mortality and PM2.5 exposure have been observed below 12 μg/m^3 (Miller et al. 2007, Eftim et al. 2008, Ostro et al. 2006, Schwartz et al. 2002 and 2008). We have reviewed the published literature, including the studies listed here, and discussed the results and their implications extensively in the report. Thus, we have revised the draft report accordingly, concluding that effects can occur at levels above 5 μg/m^3.

22. **Postponement of Staff Draft Report –** There is currently a petition for a review of CARB's earlier determination to clarify the particulate matter component of diesel exhaust as a toxic air contaminant. In view of this petition, adoption of the staff draft should be postponed pending further scientific review of recent California-specific data concerning prevalence and severity of PM-related health effects. (DTF-AS)

The relationship between long-term exposure to PM2.5 (of which diesel PM is a component) and premature death is established through epidemiological investigations and is the focus of this report. The listing of diesel PM as a toxic air contaminant is irrelevant to this review and staff's findings.

23. **PM2.5 as surrogate associated with health effects** – No threshold was detected for PM2.5 mortality effects because it is not the true causal factor; it merely co-occurs with some other compound that is responsible. (FWL)

While the possibility exists that PM2.5 is not a "causative agent" to premature mortality, the wealth of toxicological and epidemiological data to date indicates that long-term exposure to PM2.5 is a contributing factor in premature mortality. Also, how PM2.5 may cause adverse health effects, including premature mortality, is now discussed in the staff report.

24. **Source-specific PM constituents** – It is likely that each sources of constituent of PM2.5 has its own toxicity. Traffic exposure may better explain adverse health effects than unspeciated PM exposure. (FWL, DWN)

While the level of toxicity of each constituent of PM2.5 may be unique, ambient PM2.5 is a complex mixture of particles from multiple sources. The toxicity of any given mass of PM2.5 is a composite of the toxicities of all the constituents; however, PM2.5 composition is not constant on either a temporal or spatial scale at any location. While investigation of the toxicities of various source-specific PM constituents is underway, this research has not developed to the level where it can be used to inform regulatory policies.

While there have been a number of papers published recently on the effects of "traffic" on health endpoints, these papers are generally unable to separate the highly correlated effects of multiple pollutants that come from the same sources. In addition, exposure assessment in these studies is challenging. Some investigators have used such methods of exposure assessment as distance from a freeway, or vehicle counts on adjacent streets. Others have performed special purpose monitoring alongside freeways, but it is unclear how this highly localized exposure can be extrapolated to a population level. In any case, "traffic" includes PM2.5, along with several gaseous pollutants that are typically highly correlated in monitoring data that are obtained at a regional scale as well as the road-side scale. So it is unclear how attributing these effects to "traffic" helps us assess the impacts of particulate air pollution on public health.

25. **Ranges/sources of uncertainty unclear or incomplete** – Uncertainty around mortality ranges were not stated explicitly (Executive Summary). Additionally, only the uncertainty surrounding the concentration-response functions was considered; the uncertainty associated with other factors (e.g. emission estimates, PM concentrations, etc.) should be quantified. (DWN, EMA-M)

The uncertainty ranges has been included in the final version of the Executive Summary.

CARB staff did not have adequate information available to characterize the uncertainty associated with factors such as exposure estimation based on modeling or emissions data, and uniform baseline incidence rates within a county. The uncertainty section of the report gives a qualitative description of the uncertainty.

26. **Association between PM2.5 exposure and premature death not conclusively shown** – A relative risk of 2.0 is necessary to show proof of an association. Observational studies (as opposed to "randomized, controlled studies") are a weak form of supportive evidence. (JDD)

It is difficult to imagine how a randomized, controlled study of the health effects of PM2.5, or any other air pollutant, could be performed with human subjects. According to standard hypothesis testing methods, the null hypothesis in the papers in question is

that the relative risk =1, meaning that the treatment, in this case PM2.5 exposure, has no effect. If the confidence interval for the relative risk does not include 1, then the null hypothesis is rejected.

There are critics of this approach, notably John Brignell (Brignell, 2000) and Steve Milloy (www.junkscience.com), who think that the relative risk should exceed 2 to reject the null hypothesis on the ground that the standard method allows excessively high Type I error rates. This debate has arisen primarily in the context of the effects of passive smoking, and the commenter seeks to extend this debate to the subject of air pollution epidemiology. It is interesting to note that no epidemiological organization has agreed to this standard of evidence.

Confidence in a small effect estimate is not dependent only on effect size. A large sample size, and a narrow confidence interval, as in the ACS study, increases confidence in the risk estimate. In medical fields, small relative risks are frequently considered to be clinically relevant, particularly if they have narrow confidence intervals, because when the small risk is applied to a large population, there will be a noticeable difference on a population level, even if there is not a risk at the level of an individual person.

27. **Short-term exposure to PM2.5 do not cause death** – Associations between short exposure lag times and increased death rates are not plausible (JDD)

We have not asserted that exposure to PM2.5 causes death, but rather that exposure to PM2.5 is associated with risk of death. Results on short-term exposures are used only as corroborating evidence for the relationship between long-term exposure and risk of premature death. See response to comment number 5 for details.

28. **Analysis conducted in "transparent and thoughtful manner"** - Studies chosen generally were relevant to the issues at hand and to California; methodologies/how they were applied clearly explained; numerous sensitivity analyses conducted; and explicit discussion of limitations, uncertainties. (HEI)

Thank you – CARB staff worked to provide an objective, transparent treatment of the issues at hand.

29. **EPA expert elicitation process under review** – expert elicitation still in development by EPA, and will be used for PM2.5 assessment; this may help with CARB's revisions. (HEI)

The Science Advisory Board (SAB) has confirmed that the results of the U.S. EPA expert elicitation can be used to characterize the uncertainty behind the PM-mortality relationship. Further, as indicated in its findings (EPA-COUNCIL-08-002. July 2008. "Characterizing Uncertainty in Particulate Matter Benefits Using Expert Elicitation."), deriving a central estimate based on the twelve distributions is appropriate for certain applications.

The advisory can be viewed at:
http://yosemite.epa.gov/sab/sabproduct.nsf/f697818d4467059f8525724100810c37/43B91173651AED9E85257487004EA6CB/$File/EPA-COUNCIL-08-002-unsigned.pdf

30. **Sensitivity analysis could be broadened** – Multiple sensitivity analyses were conducted, which strengthens the report. However, analyses still could be broadened: expert elicitation relies on what is known at the time, so must be revisited as new findings become available. Because of this ongoing process, a wider range of possible outcomes should be included in the sensitivity analyses. (HEI)

Staff has broadened the sensitivity analysis to include Dockery et al 1993 results, as suggested by the commenter. Also, CARB staff recognizes that the PM2.5 health effects literature is continuing to evolve. Therefore, the methodology and associated sensitivity analyses will be revised as new data become available. Staff think that the sensitivity analyses used in the draft report reflect the best knowledge available at the present time, and include as wide a range of outcomes as seemed justifiable and reasonable, given the current literature.

31. **No good markers for diesel PM** – Since no good marker for ambient diesel PM has been identified, much more work is necessary before trying to quantify diesel exposure and related health effects. This topic deserves its own review, vetting, and publication in the broader literature before being incorporated into a document like this. (HEI)

We agree on the importance of an independent peer review for the diesel PM methodology. The methodology has undergone an independent peer review process by the University of California at Berkeley, Institute of the Environment. In addition, it was reviewed by Dr. Philip Hopke of Clarkson University.

While a unique chemical marker(s) has not been determined for diesel PM, various apportionment methods have nonetheless estimated the diesel contribution to ambient PM. The most specific of these methods is chemical mass balance method in which organic chemical tracer species are used. In the case of diesel exhaust, these tracer species are not unique to diesel emissions, but rather are emitted in a "signature" concentration pattern that is used to identify diesel PM emissions. Several of the assumptions in the use of this method are similar to those of the NO_X scaling method: a representative (large) set of vehicles has been sampled by ambient air filter samples, the sampled diesel PM has the same signature as the assumed (source) diesel profile, and the chemical tracer species are not transformed (react) between the source and receptor sites. The further assumption taken in the NO_X scaling method is that the sampled air has the same proportion of source emissions as that of the county as a whole (well mixed air parcel assumption); this assumption may be met by the appropriate choice of monitoring site location. Hence, we think reasonable estimates of DPM (with associated uncertainty intervals) are possible with the NO_X-scaling method.

32. **There is no set or constant relationship between engine-out PM and NO_X emissions from diesel engines**. The absolute and relative amount of PM emissions and NO_X emissions from diesel engines varies greatly depending on the application of the engine, its model and year, and the in-use duty cycle. For example, if one simply looks at emissions standards for diesel engines, there is a great deal of variability in the engine-out or vehicle-out limits for PM and NO_X. The emissions limits for NO_X are independent of the emissions limits for PM. (EMA)

We utilize an average ratio of diesel PM to total NO_X emissions, not vehicle specific diesel PM to NO_X ratio, to estimate ambient diesel PM concentrations. In this respect, the method is similar to chemical mass balance in the development of an average profile. We assume that year long averages of PM and NO_X measurements capture the average county or basin wide emission profile (for suitably selected sites).

33. **Assumptions of Cox proportional hazards model not met (See also "Epidemiologic associations (temporal)" and "Establishment of causality (confounding factors)") –** Long-term epidemiological studies of air pollution and health have used the Cox proportional hazards model, without meeting two basic assumptions: relative risks for air pollution-related health effects and confounders must be constant over time, and across different ages. (EMA-M)

The Cox proportional hazards model is a sub-class of survival models, and includes several assumptions. The model assumes that the association between PM2.5 and mortality remains constant over time. This is likely a reasonable assumption over periods of several years, although it would not be a reasonable assumption to apply over long periods of time, such as decades. Violation of this model assumption will have an increasingly important effect the longer a cohort is followed. The effect of this is demonstrated in the Enstrom (2005) paper, which followed a cohort for about 30 years, and it is beginning to appear in papers based on the Six-Cities cohort. The studies we considered all included control for known confounders (particularly smoking, as well as occupation, education, diet, and exercise) effect modifiers, and the more recent ones also included controls for spatial factors that had not been considered in earlier papers. While there is some validity to this comment when viewed in the long-term, the assumptions are reasonable over relatively short time periods (i.e., a few years), as it is unlikely that the temporal patterns of these factors would change quickly.

34. **Mischaracterization of Washington University/EPRI Veterans Cohort Study (Lipfert studies**) – Although these were cited, they were relevant to California's population and should have been more heavily weighted. Also, effects would be expected to be more pronounced in this high-risk population; the fact that none was detected point out inconsistencies in epidemiological literature. (FWL, EMA-M)

Commenters suggested that CARB staff reconsider the EPRI Veterans cohort study. The Veterans cohort studies were considered by CARB staff and the experts who participated in the U.S. EPA elicitation (See table 2a of the report). These studies

focused on male military veterans under treatment for hypertension, with 81% current or former smokers. Thus, the results do not directly apply to the general population. The CR function derived for the report was not based on any single study; rather, it was informed by this study and many others.

35. **Misuse of infant mortality studies (Woodruff et al.)** – these studies were cited in the draft report but show effects associated with PM10, not PM2.5. (FWL)

Two Woodruff studies were cited: Woodruff et al. 1997 and Woodruff et al. 2006. The 2006 study focused specifically on PM2.5 effects. The 1997 study involved PM10, and this was stated explicitly in the draft report. The size category of PM10 refers to particulate matter 10 microns or less in aerodynamic diameter and thus includes PM of 2.5 microns or less in aerodynamic diameter (i.e., PM2.5). Therefore, although it was not tested directly, it is possible that health effects in this study could have been related to PM2.5 exposure.

36. **CARB should focus more on forest fires** – Forest fires contribute to PM2.5 pollution, and CARB staff should be working to control these instead of focusing on diesel engine sources. (DWN)

Ambient PM is a mixture of all outdoor particle sources, including wood smoke, which may arise from forest fires. However, forest fires are natural events and are not a source of PM that is readily controlled. While the ARB recognizes the adverse health effects that may result from exposure to smoke from forest fires, it does not develop programs to control natural events such as forest fires. However, it does conduct research to investigate how wood smoke and forest fires might impact human health. For example, the Children's Health Study, originally funded by CARB, examined the effects of southern California wildfires on the health of children (Kunzli et al 2006, Am J Respir Crit Care Med 174:1221-1228). This study found that the children reported respiratory symptoms, such as cough and eye irritation, with smoke exposure. In addition, the likelihood of experiencing symptoms was significantly greater with six or more days of smoke exposure compared to one to five days of smoke exposure, and also with higher particulate levels, although the increase in likelihood varied among individual symptoms.

37. **Women's Health Initiative** – Co-pollutants were not given serious consideration in this study. It is unlikely that within-city risk would be so much greater than between-city risk, and the high risk levels reported are biologically implausible (EMA-M)

CARB staff agrees that the Women's Health Initiative study had strengths and weaknesses as described by some commenters. However, the CR function derived for this staff report was not based on any single study, rather it was informed by this study and many others.

It makes sense that the average PM concentration between cities would be smaller than within-city gradients. Concentrations averaged over large areas tend to mask high and

low concentrations. This observation was confirmed in the Jerrett et al. (2005) analysis of Los Angeles. Note that the potential biological mechanisms for the PM-mortality relationship is now included in the staff report.

38. **Harvard Six Cities Study: original and follow-up** – These studies did not consider confounding by co-pollutants, or other ecological covariates (e.g. socio-economic status). Laden et al. 2006 study did not show higher mortality rates in cities with high PM2.5 concentrations. Laden et al. 2006 had limited number of timepoints but was viewed favorably despite Enstrom's 2005 study being faulted for covering a limited time period. (EMA-M)

It is true that the Harvard Six Cities Follow-up study did not try to control for potential confounding by co-pollutants or other ecological covariates such as socio-economic status. With only 6 cities, such an approach is not statistically powerful enough to examine multi-pollutant models. In the original 6-cities analysis other pollutants were evaluated and fine particles and sulfates were much more strongly associated with mortality risk than other pollutants. There are many differences between the Harvard Six Cities study and Enstrom's 2005 study. The Harvard Six-Cities study, even with its weaknesses and limitations, was a much stronger prospective design, collected much better exposure data, has been analyzed, reanalyzed, and independently analyzed (Krewski et al. 2000). The biggest concern about the Enstrom 2005 study is not the limited time period but the age of the cohort after the 1979-1983 data were collected. Finally, several more recent studies (e.g., Jerrett et al. 2005) explicitly included ecological covariates in their models and still reported strong associations.