Air Pollution and Cardiovascular Disease in the California Teachers Study Cohort

FINAL REPORT

Prepared for the California Environmental Protection Agency and the California Air Resources Board
Contract # 03 - 313

Principal Investigator: Michael Lipsett, M.D.
Co-Investigators: Bart Ostro, Ph.D.; Peggy Reynolds, Ph.D.; Debbie Goldberg, MS; Andrew Hertz; Susan Hurley, MPH; Michael Jerrett, Ph.D.; Daniel Smith, Dr.P.H.

California Department of Public Health
Environmental Health Investigations Branch
850 Marina Bay Parkway, Building P, Third Floor
Richmond, California 94804

October 2007

Disclaimer

The statements and conclusions in this Report are those of the contractor and not necessarily those of the California Air Resources Board. The mention of commercial products, their source, or their use in connection with material reported herein is not to be construed as actual or implied endorsement of such products.

Acknowledgements

The authors would like to thank Ms. Cynthia Garcia, Research Division, Air Resources Board (ARB), who not only served as the Project Manager for this investigation, but also generated and provided monthly pollutant averages and interpolated pollutant surfaces for use in the exposure assessment. We would also like to acknowledge Ed Hughes, Ph.D., and Rick Burnett, Ph.D., for their able assistance with spatial autocorrelation analysis, and Ms. Theresa Saunders for assistance with administration of this contract. Finally, we would like to thank the California Teachers Study team, without whose efforts in assembling and maintaining the original cohort investigation this work would not have been possible. This Report was submitted in fulfillment of ARB Contract #03-313, Air Pollution and Cardiovascular Disease in the California Teachers Study Cohort, by the California Department of Health Services under the sponsorship of the California Air Resources Board. Work for the Draft Final Report was completed as of October 31, 2006. Revisions incorporating comments and suggestions of ARB staff and ARB's Research Screening Committee were completed May 1, 2007.

Table of Contents

	Page
Disclaimer	ii
Acknowledgements	iii
List of Tables	v
Abstract	vii
Executive Summary	viii
Background	viii
Methods	viii
Results	ix
Conclusions	x
Introduction	1
Materials and Methods	4
Study Population	4
Outcome Assessment	6
Calculation of Follow-up	7
Geocoding Study Participants' Addresses	7
Air Pollution Exposure Estimates	8
Traffic Exposure Estimates	11
Covariates	12
Statistical Analysis	14
Results	15
Discussion	47
Summary and Conclusions	55
Recommendations	60
References	63
Glossary of terms, abbreviations, and symbols	67
Appendices	68

List of Tables **page**

Table 1: Numbers of residences and total records (person-months of exposure) associated with particulate and gaseous pollutants in the California Teachers Study cohort; January 1988 – December 2002 9

Table 2. Percentage distributions of covariates at baseline among members of the CTS cohort whose addresses were geocoded 16

Table 3. California Teachers Study Participants' Counties of Residence at Baseline (1995) 19

Table 4. Descriptive statistics for air pollutants used to estimate long-term exposures among participants in the California Teachers Study 20

Table 5. Spearman correlation coefficients (r) for estimated pollutant exposures among CTS participants for the period 1995-2002 21

Table 6. Descriptive statistics for traffic, vehicle and road measures, California Teachers Study Cohort 22

Table 7. Hazard ratios for nonpollutant covariates in relation to cardiovascular mortality for CTS participants with PM2.5 data available 23

Table 8a. Results for PM2.5 using exposures from 1995-2002 for the full cohort, nonmovers and movers, disaggregated by model specification 28

Table 8b. Results for PM2.5 using exposures from 1988-2002 for the full cohort, nonmovers and movers, disaggregated by model specification 29

Table 9a. Results for PM10 using exposures from 1995-2002 for the full cohort, nonmovers and movers, disaggregated by model specification 30

Table 9b. Results for PM10 using exposures from 1988-2002 for the full cohort, nonmovers and movers, disaggregated by model specification 31

Table 10a. Results for ozone using exposures from 1995-2002 for the full cohort, nonmovers and movers, disaggregated by model specification 32

Table 10b. Results for ozone using exposures from 1988-2002 for the full cohort, nonmovers and movers, disaggregated by model specification 33

Table 11a. Hazard ratios for average air pollutant exposures (1995-2002) in relationship to disease outcomes for full CTS cohort based on pollutant interquartile ranges 35

Table 11b. Hazard ratios for average air pollutant exposures (1995-2002) in relationship to disease outcomes for nonmovers in CTS cohort based on pollutant interquartile ranges 36

Table 11c. Hazard ratios for average air pollutant exposures (1995-2002) in relationship to disease outcomes for movers in the CTS cohort based on pollutant interquartile ranges 37

Table 12a. Hazard ratios for average air pollutant exposures (1988-2002) in relationship to disease outcomes for full CTS cohort based on pollutant interquartile ranges 38

Table 12b. Hazard ratios for average air pollutant exposures (1988-2002) in relationship to disease outcomes for nonmovers in CTS cohort based on pollutant interquartile ranges 39

Table 12c. Hazard ratios for average air pollutant exposures (1988-2002) in relationship to disease outcomes for movers in the CTS cohort based on pollutant interquartile ranges 40

Table 13. Hazard ratios per interquartile range for traffic, road and vehicle metrics using full regression models, adjusted for age, race, and covariates 41

Table 14. Hazard ratios for PM2.5, per $\mu g/m^3$, using only measured (not reconstructed) data from 1999-2002 for the full cohort, nonmovers and movers 43

Table 15. Hazard ratios for cardiopulmonary (CP) mortality and AMI incidence per unit pollutant in two-pollutant models using exposures from 1995-2002 for the full cohort 44

Table 16. Hazard ratios in multi-pollutant models measuring associations between average air pollutant exposures (1999-2002) and cardiopulmonary (CP) mortality and AMI incidence among the full cohort 45

Table 17. Results using 3rd quarter ozone in single-and multi-pollutant models using exposures from 1995-2002 for the full cohort 46

Table 18. Hazard ratios for average air pollutant exposures (1995-2002) in relationship to disease outcomes for full CTS cohort based on pollutant interquartile ranges, with random effects models for spatial dependence 47

Abstract

There are few studies examining associations between long-term exposure to air pollution and adverse health outcomes. In an ongoing cohort study of over 100,000 female participants in the California Teachers Study (CTS), we developed estimates of long-term air pollution exposure at the subjects' residences and examined associations between these exposure estimates and the following health outcomes: total mortality, cardiopulmonary mortality, and incidence of both fatal and non-fatal heart attacks and stroke. In addition, we examined the potential impacts of several traffic metrics on these outcomes. In order to derive the pollutant exposure metrics, the CTS participants' addresses were linked with monthly estimates of long-term exposure to multiple air pollutants, including PM2.5 (particulate matter less than 2.5 microns in diameter), PM10, and several gases (including ozone, carbon monoxide (CO), nitrogen dioxide (NO_2), and others). The addresses were linked as well with several cross-sectional measures of traffic-related exposures from the year 2000 or later. We analyzed these relationships while adjusting for many individual-level and neighborhood variables, and undertook a variety of sensitivity analyses. We found strong and consistent associations of PM2.5 not only with total and cardiopulmonary mortality, but also with the incidence of heart attacks and stroke. We also identified somewhat less consistent relationships between one or more of these adverse outcomes and PM10, CO, NO_2 and ozone. Most of the traffic metrics were not associated with these outcomes. This study provides additional evidence that long-term exposure to air pollution is associated with mortality, and demonstrates as well that exposure to several combustion-related pollutants is associated with the incidence of new cases of heart attacks and stroke.

Executive Summary

Background

Scores of studies conducted on five continents have documented consistent associations between acute (i.e., 24-hour) exposures to ambient air pollution and daily mortality, particularly in older individuals with cardiac and respiratory diseases (California Air Resources Board 2002). In contrast, there have been only a few studies in the U.S. and Europe examining relationships between long-term exposure to air pollution and mortality. Prior studies have generally not examined whether long-term exposure to various pollutants is related to the *incidence* of cardiovascular disease, but have focused rather on *mortality* from cardiopulmonary causes. These studies cannot distinguish whether chronic exposure to air pollution, especially particulate matter (PM), plays an etiologic role in cardiovascular disease, or whether such exposures only increase the severity of pre-existing conditions. This investigation was designed not only to examine whether long-term exposures to PM and other air pollutants were associated with total and cardiopulmonary mortality, but also whether such exposures were associated with the incidence of myocardial infarction and stroke. In addition, we examined whether several traffic-related metrics were associated with the same disease outcomes.

Methods

The California Teachers Study (CTS), a large prospective cohort of active and retired female public school teachers and administrators, provided the setting for this investigation. Health outcome data were obtained through record linkage to state-wide mortality and hospitalization files for the period 1995 (the year of cohort inception) through 2002. The mortality analyses included total (non-traumatic) and cardiopulmonary (cardiovascular + pulmonary) deaths. On entry into the study, the participants completed a baseline questionnaire, which included questions on demographics, personal characteristics, and medical history, including prior myocardial infarction and stroke. Therefore, incidence analysis was limited to these two conditions. As many first occurrences of acute myocardial infarction (AMI) and stroke prove fatal, we combined both hospitalization and mortality data for each of these events in the incidence analyses.

Staff at the Air Resources Board (ARB) developed and provided us with monthly averages for PM with an aerodynamic diameter less than 2.5 μ (PM2.5), PM10, ozone, nitrogen dioxide (NO_2), nitrogen oxides (NOx), carbon monoxide (CO), and sulfur dioxide (SO_2). These monthly averages were developed for all relevant monitors operating in California from 1988 through 2002. As PM2.5 data have only been routinely collected throughout the state since 1999, however, ARB managed a separate contract to reconstruct earlier years of fine particle levels to use in this analysis. Pollutant surfaces (250m grid) of monthly average ambient concentrations were also developed by ARB staff using inverse distance weighted interpolation.

We geocoded all the participants' addresses at baseline and throughout the study period. Monthly exposure estimates for each subject's residence(s) were developed initially by spatial linkage of the geocoded locations to specific monitors and

subsequently to each monthly pollutant surface. Ultimately, we used only the interpolated pollutant surfaces for our analyses.

We also generated several traffic metrics, including distance to the nearest highway, traffic density (i.e., vehicle miles traveled within 150 and 300 m of each residence), road density (i.e., meters of roads within 150 m of each residence), and vehicle density (from 2000 Census block data). For those CTS cohort members who moved during the period 1995 through 2002, we assigned all of the (pollutant and traffic) exposure metrics to each reported address.

The statistical analysis was conducted using Cox proportional hazard regression models, adjusting for smoking status, total pack-years (for current and former smokers), body mass index, marital status, alcohol consumption, second-hand smoke exposure at home, dietary fat, fiber and calories, physical activity, menopausal status, hormone therapy use, family history of myocardial infarction and stroke, use of blood pressure medication, aspirin use, and several Census-derived contextual (neighborhood) variables (income, income inequality, education, population size, racial composition, unemployment). The main analyses covered two exposure periods: (i) 1995 through 2002, corresponding only to the follow-up period for the cohort, and (ii) 1988 through 2002, in order to examine whether a longer period of exposure might influence the results. The analysis included estimation of hazard ratios (analogous to relative risks) for the whole cohort, for those who did not move during the follow-up period ("nonmovers"), and for movers. We also undertook several sensitivity analyses, including: (i) examining the data for evidence of spatial autocorrelation; (ii) limiting the PM2.5 analysis to measured values only (1999 through 2002); (iii) running two-pollutant models for cardiopulmonary mortality and AMI incidence for the major pollutants; (iv) examining two-pollutant models for PM2.5 and ozone, with both limited only to the period 1999 through 2002; and (v) examining the impact of using values of ozone measured only during the third quarter of each year (summer).

Results

In single pollutant models, PM2.5 was associated with all-cause and cardiopulmonary mortality, as well as with incidence of AMI and stroke. The estimates for cardiopulmonary mortality and AMI incidence remained elevated and statistically significant in two-pollutant models. However, the hazard ratios using imputed PM2.5 data were of considerably greater magnitude than in virtually all other published studies, which led us to believe that the historical data probably underestimated exposures and overestimated the particle-associated risks. In contrast, in single-pollutant models using only measured data, PM2.5 was still associated with these four outcomes, with hazard ratios more concordant with other studies: for a 10 $\mu g/m^3$ increase in long-term exposure to PM2.5, the estimated hazard ratios were: all-cause mortality 1.08 (95% CI 1.00 – 1.16); cardiopulmonary mortality 1.13 (95% CI 1.02 – 1.25); AMI incidence 1.14 (95% CI 1.02 – 1.27); stroke incidence 1.14 (1.00 – 1.29). In a two-pollutant model with ozone, the PM2.5 hazard ratios remained elevated and statistically significant.

There were no significant associations between PM10 and any outcomes when the exposure period was limited to 1995 through 2002. In contrast, with the longer exposure period (1988 through 2002), PM10 exposures were significantly associated with all four

outcomes. The magnitudes of the hazard ratios were similar to those for the measured PM2.5 data for the period 1999 through 2002, and were much lower than the hazard ratios involving the historical PM2.5 data.

Single-pollutant models for ozone indicated an association with AMI incidence during the 1995 through 2002 exposure period, and with all four outcomes when the longer exposure period was used. However, with PM2.5, NO_2, or CO in the model for cardiopulmonary mortality and AMI, ozone was no longer positively associated with these outcomes, suggesting that single-pollutant ozone results may have been confounded by co-pollutants.

Due to the restrictions we placed on spatial interpolations for CO, NO_2, and SO_2, there were small numbers of participants (about 10 – 15% of the cohort) in all models involving these pollutants. Nevertheless, these traffic-associated gases were associated with all-cause, but not cardiopulmonary, mortality (for both exposure periods), and with circulatory events (NO_2 and CO with AMI and stroke incidence for both exposure periods, and NOx with AMI incidence for the longer exposure period). Even though the CO hazard ratios remained elevated and significant in two-pollutant models, as did NO_2 with ozone in the model, strong correlations involving these and other pollutants suggest that these results be interpreted with caution. Sulfur dioxide was generally not associated with any adverse outcomes.

Of the traffic metrics, road density was associated with both all-cause mortality and stroke incidence, while the results for vehicle density suggested a weaker association with stroke incidence. However, the other measures of traffic exposure showed no association with any of the outcomes.

The regression analyses in general revealed few differences in the pollutant-associated results between movers and nonmovers, except that the associations based on measured PM2.5 (1999 – 2002 only) with all four outcomes were of substantially greater magnitude among the subcohort of movers compared with nonmovers. In general, the results of this analysis did not appear to be affected by spatial autocorrelation, but we experienced difficulties in assessing this phenomenon for PM2.5.

Conclusions

This is the second largest cohort study undertaken to date examining the effects of long-term exposure to air pollution. Because we were able to link pollutant metrics with the participants' geocoded addresses, we were able to realize a degree of temporal and spatial resolution of exposure not previously attained. The results of this investigation are consistent with the published literature in demonstrating consistent associations between PM2.5 and both total and cardiopulmonary mortality. In addition, we were able to demonstrate associations of PM2.5 with incidence of AMI and stroke. The method of developing the historical PM2.5 database likely led to systematic underestimation of the variance of actual PM2.5 concentrations and therefore overestimation of the associated hazard ratios. Therefore, in our opinion, the PM2.5 hazard ratios based on measured data from 1999 – 2002 represent the most appropriate quantitative effect estimates at this time.

Long-term association with PM10 also appeared to be associated with mortality and with the incidence of AMI and stroke. Other traffic related pollutants, including CO and NO_2, were also positively associated with the adverse outcomes in this investigation,

whereas SO_2 was not. While ozone appeared initially to be associated with AMI (1995 through 2002) and with all four outcomes, using a longer exposure period (1988 through 2002), these results appear to have been confounded by exposure to other pollutants.

Except for vehicle density, the traffic metrics used in this analysis did not appear to be associated with mortality or circulatory disease incidence. While there may in fact be no relationship for these other traffic metrics, exposure misclassification or our analytical approach may have also affected the results.

There were several issues raised during the course of this study that we were unable to address, but would recommend for future research, including: (i) examination of critical windows of exposure; (ii) estimation of effects among never-smokers only; (iii) identification of specific diagnostic categories responsible for the majority of cardiopulmonary deaths; and (iv) evaluation of whether there is any effect modification of responses to air pollution by specific individual characteristics, such as chronic co-morbidities or obesity. These and other future research topics are addressed in the Recommendations section of this report.

Introduction

Scores of studies conducted on five continents have documented consistent associations between acute (i.e., 24-hour) exposures to ambient particulate matter (PM) and daily mortality, particularly in older individuals with cardiac and respiratory diseases (California Air Resources Board 2002). These findings are supported by numerous reports of cardiac and respiratory morbidity related to short-term exposure to PM, ozone, and other pollutants (U.S. Environmental Protection Agency 2004; ARB/OEHHA 2005; World Health Organization 2006). These remarkably consistent associations suggest that exposure to ambient air pollution is a risk factor for exacerbation of pre-existing cardiac and respiratory illnesses, though pathophysiological mechanisms are incompletely understood.

The evidence for a relationship between long-term exposure to air pollution and the development of cardiac or respiratory diseases, however, is relatively sparse. Several studies have examined associations between long-term exposure to air pollution and mortality: the Harvard Six-Cities Study, participants in the American Cancer Society's Cancer Prevention Studies I and II, the Adventist Health and Smog Study (AHSMOG) and the Netherlands Cohort Study on Diet and Cancer (Dockery et al. 1993; Laden et al. 2006; Pope et al. 1995, 2002; Abbey et al. 1999; Hoek et al. 2002, Enstrom 2005; Nafstad et al. 2004). All have found associations between at least one pollutant metric and one mortality category, but neither the quantitative nor the qualitative results are entirely consistent. For example, Pope et al. examined the mortality experience of over 500,000 adults in 151 U.S. cities who participated in the American Cancer Society Cancer Prevention Study II (ACS CPS II) cohort (Pope et al. 1995, 2002). After controlling for individual risk factors such as smoking, occupational exposures, body mass index, and alcohol consumption, average fine particle measurements in these metropolitan areas were associated with small, but significant, increases in relative risks (RRs) per 10 $\mu g/m^3$ increase in PM2.5 for all-cause (1.06, 95 % CI = 1.02-1.11), cardiopulmonary (1.09, 95 % CI = 1.03-1.16), and lung cancer (1.14, 95 % CI = 1.04-1.23) mortality (Pope et al. 2002). Although the Harvard study included far fewer subjects (n=8,111 at baseline), those investigators found similar results for several Midwestern and Eastern cities (Dockery et al. 1993), which was confirmed in a recent follow-up analysis suggesting that decreases in PM2.5 were associated with decreased relative risks for all three mortality categories (Laden et al. 2006). For the follow-up period of 1974 – 1998, these investigators reported relative risks per 10 $\mu g/m^3$ increase in PM2.5 of 1.16 (95 % CI = 1.07-1.26) for all-cause, 1.28 (95 % CI = 1.28-1.44) for cardiovascular, and 1.27 (95 % CI = 0.96-0.69) for lung cancer mortality. The Harvard and ACS CPS II studies used only one air pollution monitoring site per city, though the areas of coverage for these monitors were significantly different. The Harvard study deployed monitors specifically for the purpose of the study and had relatively small spatial catchment areas, while the ACS CPS II monitors covered very large areas, often larger than Metropolitan Statistical Areas.

In contrast, three California-specific studies have reported fewer associations and less consistent results. Enstrom (2005) found essentially no relationship between fine PM and all-cause mortality among nearly 50,000 elderly California participants in Cancer Prevention Study I (ACS CPS I) over the period 1973-2002. However, in this study,

assignment of exposure was relatively crude, and therefore measurement error may have played a significant role in these results. Each individual was assigned county-wide PM2.5 levels measured only during the period 1979-1983. In addition, the ACS CPS I cohort was much older at intake and may have had different indoor/outdoor exposure-related behaviors relative to other cohorts in air pollution studies. Interestingly, for the initial follow-up period (1973-82), which included several years of fine particle measurements, Enstrom reported a RR for all-cause mortality of 1.04 (95%CI 1.01-1.07) per 10 $\mu g/m^3$ increase in PM2.5.

In the AHSMOG study (N = 6,338 Seventh Day Adventists throughout California), Abbey and colleagues (1999) attempted to reduce exposure measurement error by interpolating pollutant monitoring station data to the zip code centroids for the participants' home and work addresses. Those investigators found associations of long-term-exposure to particulate matter and ozone with deaths related to diseases of the lung, but, unlike the Harvard and ACS II studies, not with those involving the cardiovascular system. More recently, however, Chen et al. (2005) analyzed fatal coronary heart disease (CHD) events in a subset (n=3,239) of the AHSMOG cohort who had been followed for 22 years. These investigators reported associations of PM10, PM2.5, and coarse particles (i.e., PM10-PM2.5) with fatal CHD in women, but not in men. In a multivariate model that was also adjusted for ozone, the RR for fatal CHD per 10 $\mu g/m^3$ increase in PM2.5 was 2.00 (95%CI 1.51-2.64).

In a recent re-analysis of the ACS CPS II data for 22,905 residents of the Los Angeles basin, Jerrett et al. (2005) used multiple kriging algorithms to interpolate PM2.5 and ozone data to the 267 zip codes encompassing the subjects' residential addresses. These investigators reported increased relative risks associated with a 10 $\mu g/m^3$ increase in PM2.5 of 1.17 (95% CI 1.05-1.30) for all-cause mortality, 1.39 (95% CI 1.12-1.73) for deaths from ischemic heart disease, and 1.44 (95% CI 0.98-2.11) for lung cancer, controlling for 44 individual-level covariates. The estimates were robust to the inclusion of ozone and traffic metrics in the models, but were reduced with the addition of social contextual variables, such as neighborhood income, education, and racial composition.

Three European cohort studies have also investigated the impact of chronic exposure to air pollution on mortality. Nafstad et al. (2004) examined the mortality experience of 16, 209 middle-age Norwegian men followed from 1972-73 through 1998. Their exposure assessment consisted of a complex modeled GIS estimation of long-term nitrogen oxide and sulfur dioxide concentrations at the participants' residential addresses. PM was not considered in this analysis because PM measurement methods changed during the period of observation. Nafstad et al. (2004) reported significantly increased relative risks associated with nitrogen oxides, but not sulfur dioxide, for all-cause mortality, as well as deaths from lung cancer, nonmalignant respiratory disease, and respiratory disease. In a study of 5,000 adults followed from 1986 to 94 in the Netherlands, Hoek et al. (2002) found strong associations between several pollutant metrics (black smoke, nitrogen dioxide, and living near a major road) and cardiopulmonary mortality, with somewhat lower risks for all-cause mortality. Living near a major road was the exposure variable most strongly associated with cardiopulmonary mortality (RR = 1.95, 95% CI, 1.09-3.52), adjusting for regional black smoke, age, sex, education, body mass, occupation, exposure to tobacco smoke, and a composite of neighborhood socioeconomic factors. Along with the investigation by

Jerrett et al. (2005), this study suggests that intra-urban differences in chronic exposure to air pollution, particularly those associated with traffic emissions, may be at least as great as those associated with mean inter-urban differences.

In considering potential mechanisms for disease development, there are common threads linking air pollution exposures with the above outcomes. Several common pollutants, such as ozone and PM (including polycyclic aromatic hydrocarbons or PAHs), are recognized to induce oxidative stress in the lung, which may also result in both local and systemic pro-inflammatory effects (Li et al. 2003; Gurgueira et al. 2002). Asthma and COPD have both long been understood to involve chronic airway inflammation, while the importance of inflammatory processes in the etiology of cardiovascular disease has only recently been recognized (Brook et al 2004; Tousoulis et al. 2003). Such processes may underlie associations between cardiovascular morbidity or mortality and air pollution (Dhalla et al., 2000; Donaldson et al., 2001). Electrophiles in PM, including PAHs, are thought not only to be capable of inducing intracellular cascades resulting in the synthesis of pro-inflammatory cytokines, but also to produce mutations in key genes related to carcinogenesis, such as those involved with DNA repair and tumor suppression. The numerous oxidants and electrophilic compounds in ambient air pollution provide a basis for inferring plausible pathophysiological mechanisms.

The extent to which long-term exposure to particulate matter, ozone or any other air pollutant may be linked with cardiac, respiratory, or malignant disease is an issue of enormous public health and regulatory significance. State and federal annual average ambient air quality standards for particulate matter are based primarily on the results of the two largest cohort studies (Dockery et al. 1993; Pope et al. 2002). However, there are inconsistencies among the published studies with respect to the magnitudes of effect associated with different pollution metrics, which may be related in part to exposure misclassification, as noted above. Nevertheless, other major "lifestyle" differences between Seventh Day Adventist participants in the AHSMOG cohort and the general population may limit the extent to which the results of this study can be generalized.

Cardiovascular disease is by far the largest cause of mortality in the U.S. population. With one very recent exception, prior studies have not examined whether long-term exposure to various pollutants is related to the incidence of cardiovascular disease, but have focused only on mortality from cardiopulmonary causes. Therefore, these studies cannot distinguish whether chronic exposure to pollution plays an etiologic role in cardiovascular disease, or whether such exposures only increase the severity of pre-existing conditions.

This investigation was designed not only to examine whether long-term exposures to PM2.5, PM10, and several gaseous air pollutants were associated with total and cardiopulmonary mortality, but also whether such exposures were associated with the incidence of myocardial infarction and stroke. We proposed to undertake this investigation by adding an air pollution component to a cohort study of over 100,000 female teachers in California. This cohort offered a unique opportunity to examine the relationships between specific air pollutants and chronic disease outcomes more carefully than had been done previously. The prevalence of active smoking in this cohort was quite low (about five percent at baseline), allowing for careful examination of the impact of air pollution exposures. In addition, because of the similarity of the educational backgrounds and working environments for the cohort members, significant confounding

or effect modification by these factors is unlikely. The sheer size of the study also allowed for markedly greater statistical power than any of the earlier cohort studies except for the national ACS CPS II. The vast majority of the cohort continues to reside in California, where large metropolitan areas contain an unparalleled air pollution monitoring network, providing opportunities for refined exposure assessment over extended periods of time, which has not been feasible in the other cohorts. The participants' home addresses were previously geocoded, which also allowed for analyses of the impacts of exposure to local traffic emissions. In brief, there were a variety of advantages that this dataset would offer over prior investigations.

Materials and Methods

Study Population

The California Teachers' Study (CTS) is a prospective study of 133,479 current and former female public school employees who completed baseline questionnaires in response to two mailings to all 329,684 active and retired female enrollees in the State Teachers Retirement System (STRS). The STRS is a quasi-public agency that manages retirement investments for California educators (teachers and administrators) employed in public school systems, including all primary and secondary school teachers as well as faculty in the state junior college and university systems. STRS members are employed in approximately 1,160 public school districts, community college districts, county offices of education, and state reporting entities throughout California. All California public school employees must pay into and receive retirement benefits through STRS; membership continues as long as retirement contributions remain on deposit with the program. The STRS maintains current address information on members even after they retire or leave California.

The CTS cohort was established in 1995 using State of California cigarette tax revenues, initially for investigating a previously reported excess incidence of breast cancer in California teachers. The study was developed by a consortium of investigators from the California Department of Health Services and three active research institutions that manage regional registry operations as part of California's statewide cancer surveillance program (the Northern California Cancer Center, the University of California, Irvine and the University of Southern California (USC)). The design and on-going follow-up of the CTS cohort is a collaborative effort of the study's co-investigative group representing researchers with diverse and complementary areas of expertise. One of the co-investigators for this study, Dr. Peggy Reynolds, is a founding member of the CTS and remains an active member of its Steering Committee.

There have been four waves of questionnaires for the CTS: 1995 (baseline or Wave I), Wave II (1997), Wave III (2000), and Wave IV (2005). Self-reported chronic conditions were recorded in Wave I, and hospitalization information was also collected in Waves II and III. For this investigation only the responses to Waves I through III were utilized. In these analyses, survey data from these questionnaires were used to characterize factors that may be important confounders/effect modifiers of the air pollution/health outcome relationships. Data on numerous other potential risk factors for chronic disease were included in the CTS database, including (among others) relevant

demographics such as age and race, exercise patterns, diet, active and passive smoking exposures, alcohol use, weight, individual and family medical histories, and use of medications and hormones. The baseline questionnaire also included questions on history of chronic disease, including specifically any history of a prior myocardial infarction or stroke.

The CTS cohort is well characterized, diverse, and represents a range of socioeconomic levels, depending in part on spousal income. Participants' ages varied from 20 to over 100 years at enrollment, with a mean of 54 years. The cohort is multiethnic but primarily white (86.7%) and born in the United States (93.6%). At baseline, 124,514 (93.3%) of the women lived in California. Approximately 78% of the cohort members were elementary or high school teachers for the majority of their careers and over 50% were employed in the school system more than 15 years. A full description of the CTS cohort is available elsewhere (Bernstein et al. 2002).

Record linkage is conducted annually to mortality files and to the statewide cancer registry (both administered by the California Department of Health Services), and to a statewide file of hospitalization data, administered by the Office of Statewide Health Planning and Development (OSHPD). Ongoing routine follow-up of the cohort includes updating name and residential information of CTS members for the purposes of future contacts as well as for outcome linkage. The primary method for address updates comes via the US Postal Service (USPS). Of the approximately 86,000 name and address changes recorded for the cohort through January 2002, 60% came via notification of changes of address made to the USPS. In preparation for each of the nonprofit mailings sent to CTS members, the address data files are processed electronically by a USPS-designated service agency. The second largest source of ongoing name and address updates is the cohort members themselves -- via changes of address recorded on questionnaire covers, postage paid postcards included in annual newsletter mailings, telephone calls to a 24-hour toll free line, and e-mail notifications. An additional form of active follow-up involves periodic phoning of cohort members who have not responded to mailings. Projects completed by skilled medical interviewers as well as high-volume outbound call centers facilitate collection of additional address change and address verification information.

Supplementing these "active" follow-up methods, "passive" methods are also extensively utilized for the purpose of verification of state of residence or vital status. Since the main outcome measures for the cohort involve record linkage against statewide cancer registry and hospital discharge databases, confirmation of residency in California is critical. Resources such as drivers license records, voter registration rolls, property tax files, and Social Security vital status records are used. All these resources added to the active follow-up contribute to a "cohort viability score," which is a composite measure of the various forms of residency confirmation. This score shows that slightly more than 95% of the cohort had verifications in 2000 or later.

Use of data on human subjects in the main CTS cohort study was reviewed and initially approved by the California Committee for the Protection of Human Subjects, Health and Human Services Agency, in June 1995 and annually thereafter. The same committee approved use of the CTS data specifically for this investigation in August 2004 and renewed the approval in 2005 and 2006.

Outcome Assessment

There are several sources of information on health outcomes among the CTS cohort. As noted above, record linkages of the CTS cohort were conducted annually through 2002 to mortality and hospital discharge data by CTS co-investigators at USC. Mortality outcomes were ascertained through files administered by the California Department of Health Services as well as with the Social Security Administration (SSA) death master file. These linkages were performed using probabilistic record linkage utilizing AUTOMATCH (Jaro, 1995). Secure internet-based retrieval software permits real-time viewing and printing of California death certificates. Mortality data through December 31, 2002, were utilized for this study's total mortality (excluding external causes) and cardiopulmonary (CP) (cardiovascular + pulmonary) mortality outcomes. Mortality ICD-9 codes were used for deaths occurring in 1995 through 1998 and ICD-10 codes were used for deaths during 1999 through 2002. Total mortality (excluding external causes) included all ICD-9 codes except for those greater than 800 and all ICD-10 codes except S,T,U,V,W,X,Y, and Z. Cardiopulmonary mortality included ICD-9 codes 390-459 and 460-519 and ICD-10 codes I00-I99 and J00-J98.

Incidence data were ascertained through linkage with hospital discharge data collected and maintained by OSHPD. This database contains information about admissions to all California hospitals except military facilities. The data include up to 25 diagnoses and up to 21 procedures per admission. OSHPD does not collect the name of the patient, but since 1991 this database has included collected Social Security number, date of birth, sex, race, and ethnicity. Using these variables, probabilistic record linkages are performed annually under separate funding (NCI R01 CA77398). The record linkage used in this study was conducted through December 31, 2002, in order to determine the incidence of acute myocardial infarction (AMI) (ICD-9 codes 410 and 414) and stroke (ICD-9 codes 431-434 and 436). Women were excluded from the AMI incidence analysis if they reported a previous occurrence of heart attack or myocardial infarction on the baseline questionnaire or had a hospitalization due to an AMI prior to completing their baseline questionnaire. Women were excluded from the stroke incidence analysis if they reported a previous occurrence of stroke on the baseline questionnaire or had a hospitalization due to a stroke prior to completing their baseline questionnaire. The initial episode of AMI or stroke is often fatal; therefore, in order to capture incidence of these events, we created a variable combining hospitalization and death for each of these outcomes. In the incidence analyses, unique subject identifiers allowed us to avoid double-counting a hospitalization for either of these events that subsequently resulted in a fatality.

Calculation of Follow-up

Person-days at risk were based on the first seven years of follow up. For women who remained in California for the entire follow-up period, we calculated person-days at risk as the number of days between the date each woman joined the cohort (i.e., the date she completed her baseline questionnaire) and the earliest of three dates: the date of death (for mortality analyses) or the date of hospitalization or death (for AMI or stroke incidence analyses), or December 31, 2002. For women who moved out of state, person-days at risk were calculated from enrollment until the date of her first non-California address, assuming that she had not been hospitalized for AMI or stroke before then. In the analysis of incidence of myocardial infarction and stroke, we only counted until the first episode: time after a first hospitalization for either of these conditions was censored. Women who died from some cause other than the outcome of interest during the follow-up period were censored at the time of their death.

Geocoding Study Participants' Addresses

The baseline addresses of the CTS participants had been previously geocoded: we initially intended to geocode only new addresses for individuals relocating within California during the period 1995 through 2002. However, we were concerned that the baseline geocoding, which had been done several years earlier, would not be of equivalent precision or quality to the geocoding for post-1995 addresses. Moreover, we wanted the source data from which we were geocoding to be consistent with the source data from which we were estimating traffic exposure. Therefore, we decided to re-geocode the baseline along with the post-baseline addresses. We received 207,110 address records for period 1995 through 2002 from the CTS data center (at the University of Southern California). Each record represents a name change, move, or residency confirmation supplied from various sources, including the baseline and follow-up study questionnaires, respondent telephone calls or correspondence, and linkages to DMV, postal, and other records. Prior to geocoding the address information, we reviewed the data, and in consultation with USC staff, eliminated duplicate records, as well as records that were likely to be address corrections or reformatting rather than true moves. Furthermore, we restricted the file to those CTS members who resided in California at the time of the baseline questionnaire. After removal of the duplicate addresses, name change records, and non-California addresses (7,981 moved out of CA during this period), 173,071 address records (for 124,614 individuals) remained to be geocoded. This total included 88,666 single addresses of members who had never moved, with the balance representing multiple addresses of 35,948 movers, some of whom relocated more than once.

These address records were standardized to USPS format using ZP4 address correction software (Version 58, Semaphore Corporation, Pismo Beach, CA). Post Office Box addresses (N = 9,476) were flagged as non-geocodeable, leaving 163,595 addresses as potentially geocodeable. These addresses were batch geocoded against three different street datasets: Navigational Technologies (2005q2), Geographic Data Technology's Dynamap2000 (2005q1), and TeleAtlas (2004q1). A total of 150,105 addresses were successfully geocoded via batch processing, with success defined as a

match score of 100 on all three street datasets. However, to maintain consistency with the traffic count data, we created the exposure database using the geocoding coordinates from Dynamap 2000. Manual review of the remaining 13,490 addresses resulted in an additional 13,073 residences being geocoded with sufficient precision for point scale analyses, again using Dynamap 2000 as the default source of geocoding coordinates. Of the total numbers of address records, approximately 99% of those that were not P.O. boxes were geocoded, though among the movers, this figure was approximately 97%. Thus, a total of 163,178 residences were available for estimating exposure to air pollutants.

Air Pollution Exposure Estimates

We planned to undertake several different approaches to estimate long-term pollutant exposures involving three exposure periods: (1) from cohort inception in 1995 through 2002; (2) from January 1988 (the earliest time that reasonably complete statewide data were available for PM10 as well as the gaseous pollutants) until 1995; and (3) the entire period from January 1988 through December 2002. ARB staff provided us with monthly average concentrations from fixed monitoring sites for PM10, ozone, nitrogen dioxide (NO_2), carbon monoxide (CO), sulfur dioxide (SO_2), and nitrogen oxides (NOx) from January 1988 through December 2002. Subsequently, ARB staff developed pollutant surfaces for these pollutants using inverse distance weighting (Appendix 1).

Routine ambient air monitoring for $PM_{2.5}$ were available from 1999 through 2002. In order to utilize the full health outcome dataset in the analyses, however, an historical PM2.5 database extending at least from the beginning of 1988 was needed. In order to generate this PM2.5 database, the ARB funded and managed a separate contract with Dr. Charles Blanchard (Blanchard and Tanenbaum 2005). In that project, the contractor combined data from different monitoring programs, accounting for differences in measurement methods and accuracy, to create a database of estimated monthly-average fine PM mass concentrations and their uncertainties for the period from 1980 through 2002. Dr. Blanchard reconstructed FRM-equivalent PM2.5 mass concentrations from several other types of measurements, establishing conversion factors to standardize fine mass measurements from other networks to FRM equivalents. The other monitoring sites and samplers included the ARB dichotomous sampler network and several special studies conducted prior to implementation of the FRM network. Where alternative measurements of fine PM mass were not available, he reconstructed fine mass from PM components measured in other size fractions, light absorption, or light scattering (Blanchard and Tanenbaum 2005). A summary explanation of Dr. Blanchard's work is provided in Appendix 2.

We collaborated with ARB staff in validating monitor locations and produced the spatial file that we and ARB used jointly in developing the exposure estimates. Specifically, we had received from ARB staff a file with monitor addresses and geocoded coordinates. In addition, the ARB website listed GPS coordinates for a large number of monitors. We re-geocoded the monitor addresses in the file from ARB and compared these coordinates with those already in the file on the ARB website. There were a

number of non-matching coordinates that ultimately were addressed by comparing them with aerial photographs to identify the most likely locations.

Our initial approach to assigning exposure was to link the teachers' residences with the nearest monitoring stations, starting with ozone. This involved assigning average ozone exposure values for each month of the follow-up time period to the geocoded teachers' residences. From 1995 on, it was possible to calculate monthly exposure values for both movers and nonmovers. For each cohort member who moved, an additional analysis was performed using estimates of pollutant exposure based on the person-months at each residence. However, for the period from January 1988 until 1995, we did not have residential histories and therefore assumed, for purposes of assigning exposures, that the subjects had resided at their baseline (1995) addresses during that period.

Monitoring stations in California are (with few exceptions) assigned one of five spatial scale designations meant to give a rough approximation of the radial distance over which the monitor readings would be representative: micro-scale, middle, neighborhood, urban and regional. In consultation with ARB staff, we decided to exclude from this analysis all monitors designated as micro- or middle scale, as these were considered to be representative of ambient concentrations only up to about 0.5 km. However, some monitors in the ARB database have no scale designation. We conducted analyses to determine how many teachers would likely be affected by assigning a spatial scale (i.e., neighborhood, urban or regional) to monitors missing a spatial scale designation. We found that the potential benefits would be minimal, particularly since this approach would require validation with other monitors whose scale designation was already known. Therefore, these undesignated monitors were not used in the analysis. We also examined the potential benefit, in terms of the numbers of additional subjects who could be included in the analysis, of using middle-scale monitors and determined that the impact would be negligible.

Table 1: Numbers of residences and total records (person-months of exposure) associated with particulate and gaseous pollutants in the California Teachers Study cohort; January 1988 – December 2002

Pollutant	Spatial Scale (km)		Residences (in range)	Total Records
	Neighborhood	Urban/Regional		
Ozone (1-hr max)	20	50	159,914	28,784,520
PM2.5*	20	20	146,888	26,439,840
PM10	10	20	103,403	18,612,540
NO_2 (1hr max)	3	5	24,723	4,450,140
NOx (1hr max)	3	5	24,723	4,450,140
CO (8-hr avg max)	3	5	16,291	2,932,380
SO_2 (24hr avg)	3	5	33,045	5,948,100

*There was no spatial scale associated with the PM2.5 monitors. We elected a 20 km buffer in consultation with ARB staff.

We created buffers around the remaining monitors based on the latters' spatial scales (Table 1). Subjects whose residences were not within the set of buffers were excluded from the analyses. Monitors with no measurement data during the period of interest (1988 through 2002) were also excluded. Starting with ozone, we calculated distances from each teacher's residence to all monitors within range (based on spatial scale), then assigned the nearest monitor to each residence and linked monitor measurements to the teachers' residences. Table 1 lists the numbers of residences and records associated with the different scale designations for each pollutant.

Our initial analysis revealed that a considerable number of "nearest" monitors had incomplete temporal coverage, which created significant issues of missing data when we linked the monthly average values with the teachers' residences. We then calculated a new estimate of exposure by averaging measurements for all monitors whose "range of representativeness" (based on spatial scale designation) encompassed a given residence. This resulted in far fewer person-months with missing ozone estimates.

Pollutant surfaces (250 m grid) of monthly average ambient concentrations for O_3, PM_{10}, NO_2, NO_X, SO_2, and CO were developed by ARB staff using inverse distance weighted (IDW) interpolation (Appendix 1). To maximize the spatial coverage, all available monitors were used for each month of data. However, as some monitors were added to the network or were taken offline, the numbers of monitors used to estimate the pollutant surfaces varied. ARB staff also provided interpolated average monthly PM2.5 concentration surfaces based on the dataset provided by Dr. Charles Blanchard (Blanchard and Tanenbaum 2005). Exposure estimates for each subject's residence were made by spatial linkage of the geocoded location to each monthly pollutant surface. Illustrative ArcGIS algorithms involved in creating these linkages are depicted in Appendix 3.

The initial regression analyses (see below) were undertaken using the exposure metric based on monthly pollutant values derived from the average concentrations of all monitors within range of a given residence. However, as this approach appeared to represent a crude approximation of the more formal IDW interpolation provided by ARB, we compared the output from both approaches, and found that there was little difference. Therefore, all subsequent analyses were based on exposure estimates derived from the IDW-interpolated surfaces.

The initial PM2.5 dataset had to be recreated after errors were identified in the historical data that had been provided to ARB and then to us. Specifically, Version 2 of the California PM2.5 database failed to convert two-week measurements from eight samplers to FRM-equivalent concentrations. For all site-months for which the best estimate of the monthly-average PM2.5 concentrations had been obtained from two-week sampler data, the prior estimates had to be corrected with a multiplier of 1.184 to obtain the equivalent FRM value. In addition, for a small subset of monitors, the sign of the longitude value was erroneous (positive instead of negative). After all of these corrections were made to the monitoring data, ARB re-ran the PM2.5 interpolations, which we then re-linked to the subjects' residential addresses.

Finally, we had initially intended to examine two averaging times for the gaseous pollutants as a basis for calculating person-months of exposure. As a preliminary step in this process, we evaluated the correlations between the measured values for these different averaging times for ozone (1 hr vs. 8 hr), CO (8 hr vs. 24 hr), NO_2 (1 hr vs. 24

hr), and NOx (1 hr vs. 24 hr). Each pair of pollutant metrics was very highly inter-correlated according to both Pearson and Spearman Rank methods (correlation coefficients > 0.90). Therefore, after consulting with and obtaining approval from ARB staff, we only evaluated associations of the health outcomes with person-time of exposure based on just one averaging time per pollutant.

Traffic Exposure Estimates

Several vehicular-related exposure metrics were assigned to each geocoded residence: (1) 2000 Census Block Group vehicle density; (2) proximity to a major highway (within 20 km); (3) road density (meters of road within 150 m); and (4) traffic density (vehicle miles traveled within 150 m and 300 m). All road based metrics were calculated using TeleAtlas' Dynamap 2000 (2nd quarter, 2005 release) street database.

Vehicle density was calculated using the aggregate number of vehicles in occupied housing units (variable name H046001) divided by the land area of the block group within which the teacher resided.

Distance to the nearest major highway (within 20 km) was calculated for each geocoded residence. Major highways were defined as those having a functional classification code of "A1" (primary road with limited access, e.g., an interstate or other freeway with on-ramps and off-ramps) or "A2" (primary road without limited access).

Road "density" was calculated by summing the lengths of all roads within 150 m of each geocoded residence. Since we only used one radius, the area was the same for all residences. Therefore, the value of this variable consisted of total road length per 22,500 m^2.

Traffic density was calculated by summing vehicle distance traveled on all measured roads within 150 m of a geocoded residence. Vehicle distance traveled for each road is the number of vehicles on the road multiplied by the length of the road (See Figure 1).

The Federal Highway Performance Monitoring System (HPMS) provides vehicle counts for roads. The original CalTrans street dataset that linked to HPMS data was digitized from 1:100,000 scale U.S.G.S. Digital Line Graph (DLG) maps, so the spatial precision is poor relative to current street files used for geocoding. CalTrans and GDT (now TeleAtlas) have been collaborating to conflate the CalTrans street attribute data to GDT's Dynamap street data. The conflation process is not complete and some problems were observed in the dataset. Data for Santa Cruz County was corrupt and unreadable. In the original CalTrans street dataset, all streets were geographically represented by a single street centerline. The GDT street data represent divided roads (including most freeways) by two parallel centerlines, one for each direction of travel. Traffic count data is a measure of two-way traffic on a street, so when linked to the conflated GDT streets, divided roads will be double counted. Based on the functional classification coding in the Dynamap data, divided roads ("A" followed by "1"-"4" followed by "5"-"8") were selected and the vehicle counts on those segments were halved (Figure 4). Some (approximately 1/8th) of the identifiers used to link the CalTrans streets to the HPMS traffic data appear to be different in the conflated GDT data, preventing linkage. For streets whose names were coded as a route identifier, a new ID field could be constructed

by replacing the last six numbers from the Segment ID field with the first six numbers from the street name field.

For all road-based measures, two options were considered to account for missing data: one in which missing data were excluded, and one in which missing data were assigned a non-null minimum exposure value (50 km for proximity to major highways and the minimum calculated values for road and traffic densities).

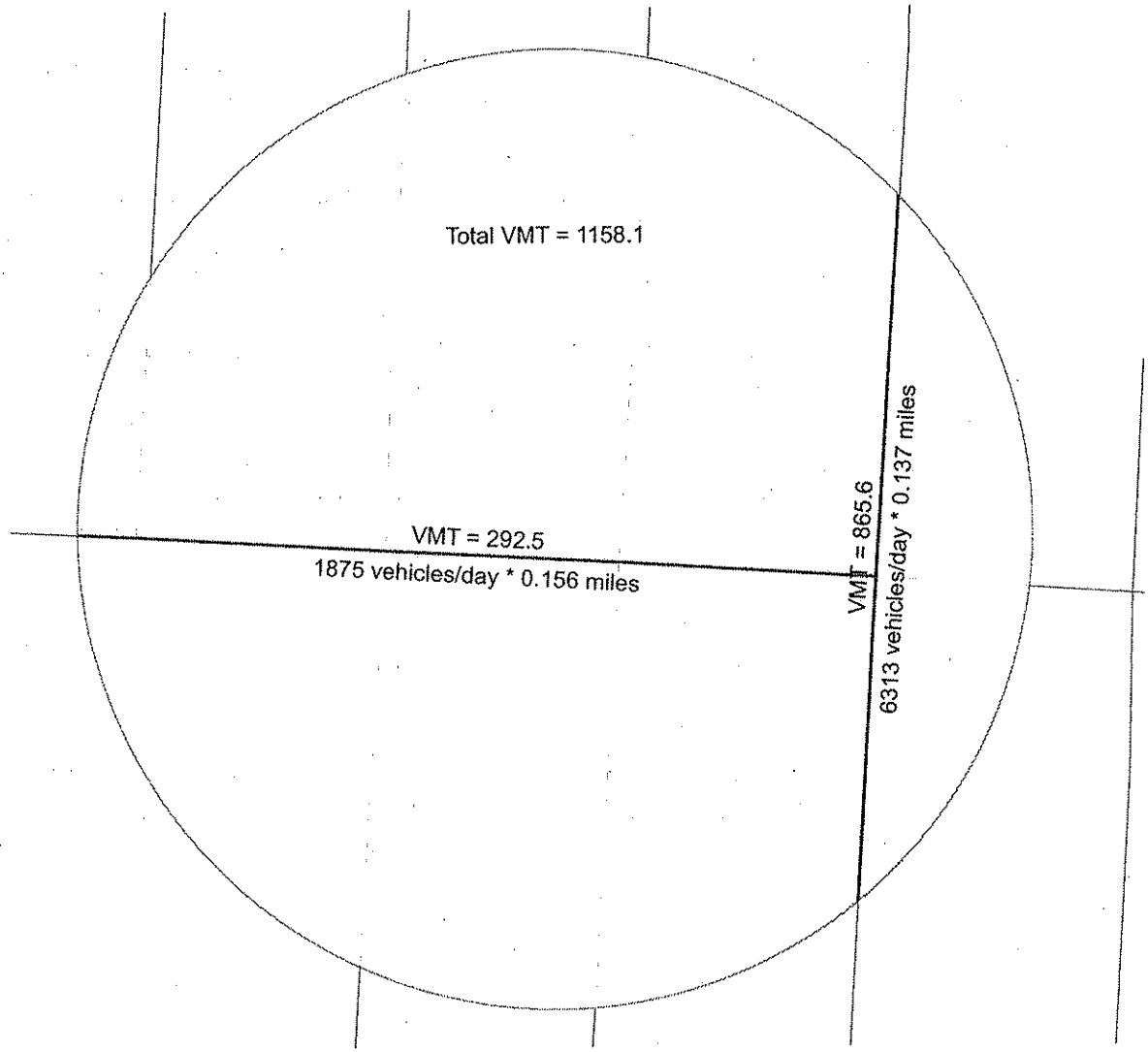


Figure 1: Example of traffic density estimation within a circular buffer zone

Covariates

Cardiovascular disease outcomes constituted the principal outcomes of interest in this analysis. Therefore; we selected most of the individual-level predictor variables for the regression analysis based on previous studies of cardiovascular disease, including investigations that examined the influence of air pollutants on cardiovascular and cardiopulmonary mortality (Dockery et al. 1993; Pope et al. 2002; Jerrett et al. 2005).

Age was divided into two-year categories between ages 30 and 69 and three-year categories between ages 70 and 89, and one category for women aged 90 and older. Race/ethnicity was categorized into three groups: non-Hispanic white, all others and unknown/not provided. Marital status categories included married/living with partner, not married (i.e., divorced, separated, widowed, never married), and unknown/missing. We based smoking status on two questions from the baseline questionnaire. Women were asked if they had ever smoked 100 or more cigarettes during their lifetime and, if so, when they started and stopped smoking. Using this information we categorized respondents as never, former, or current smokers. We measured smoking pack-years (i.e., the number of packs smoked per day times number of years smoked) as one continuous value for former and current smokers. Second-hand smoke exposures were categorized into three groups: those with exposure to household second-hand smoke, those without such exposure, and those with unknown exposure. Household second-hand smoke exposure was based on the women's report of ever having lived with a smoker.

We calculated body mass index (BMI or weight/height^2) for each participant based on her questionnaire responses regarding her weight and height. The women were grouped into BMI categories as follows: less than 18.5 kg/m^2, 18.5-24.9 kg/m^2, 25-29.9 kg/m^2, 30-39.9 kg/m^2, 40 kg/m^2 or more, and height or weight not provided. Physical activity, defined as the average number of hours per week of moderate or strenuous activity over a women's lifetime, was categorized as low (less than two hours per week), medium (2-4.90 hours per week), high (4.91 hours per week or more), and not provided. Alcohol consumption categories included non-drinkers, separate dummy variables for any beer, wine and alcohol consumption, and unknown/missing. We also created tertiles of daily caloric intake (less than 1281.63 kcal, 1281.63-1718.57 kcal, 1718.58 kcal or more, and unknown), fat (less than 41.80 g/day, 41.80-62.74 g/day, 62.75 g/day or more, and unknown) and fiber (less than 11.08 g/day, 11.08-16.01 g/day, 16.02 g/day or more, and unknown).

Women's menopausal status was defined as pre-/perimenopausal, post-menopausal, and not able to determine. We categorized women's hormone therapy use as never used estrogens, used estrogens, and unable to determine.

Family history of AMI or stroke was defined as an occurrence of these events in either the respondent's mother or father. These were then summarized into dichotomous variables. High blood pressure medication and aspirin use were combined and summarized into categories including no medication, low, medium, and high dosages.

Ecological variables at the block group level were derived from 1990 Census data in order to control for "contextual" neighborhood confounding (Jerrett et al. 2005). These contextual variables represent social, economic, and environmental settings at a group level that may be associated with disease outcomes at the individual level. Such effects may interact with individual-level variables or may be independently associated with the outcomes. The contextual effects in this study were selected based on their identification in prior studies, particularly the ACS CPS II studies (Jerrett et al. 2005), and include ethnic/racial composition (black, white, and Hispanic), income, unemployment, population size, income inequality, and education.

Statistical Analysis

The statistical analyses in this study were limited to cohort members who were living in California and at least 30 years old at the time they completed their baseline questionnaire, who were successfully geocoded, and had available information on all variables used in the statistical models. Because the U.S. Census Bureau suppresses data on block groups with very small population counts, some contextual variables were not available for all block groups. Therefore, our analyses were limited to respondents for whom this information was available.

Analyses of incident AMI and stroke were further restricted to those women who reported no history of such events on the baseline questionnaire and had no prior outcome-specific occurrence reported in the OSHPD database. As many first occurrences of AMI and stroke prove fatal, we included both hospitalization and mortality data in the analyses of incidence of these outcomes. Unique subject identification codes allow for incidence analyses combining hospital admission and mortality data without "double-counting" any events that result in both hospitalization and death. These were based on the primary cause of death or principal admission diagnosis; however, we also undertook sensitivity analyses involving up to two underlying cause(s) of death (if any) specified on the death certificate.

We used Cox proportional hazards models to estimate hazard rate ratios (HRs) and 95% confidence intervals (CI) for each pollutant and traffic metric of interest. HRs and 95% CIs were also scaled to the interquartile range (IQR), based on distributions for non-event women in each particular pollutant and traffic analysis. Initial models were adjusted for age and race strata only. Subsequent models were further adjusted for the personal risk factors (marital status, smoking status, smoking pack-years, second-hand smoke exposure, BMI, physical activity, alcohol consumption, dietary caloric intake, fat, and fiber consumption), female-specific risk factors (menopausal status and hormone therapy), and contextual variables (racial composition, income, unemployment, population size, income inequality, and education). Models assessing cardiopulmonary mortality and AMI and stroke incidence were also adjusted for additional cardiovascular risk factors (family history of AMI or stroke, high blood pressure medication, and aspirin use).

It is possible that disease incidence and survival times may be more similar among study subjects who live in communities closer together than among those who live in communities separated by greater distances, regardless of air pollution exposures. In this instance, the data may be subject to spatial autocorrelation. To the extent that the latter is related to missing variables or systematic misclassification of exposure to risk factors spatially associated with the air pollutant variables, the risk estimates of pollutant-associated disease incidence or mortality may be biased. To address this issue, we undertook sensitivity analyses using a spatial random-effects model similar to that employed in an analysis of the ACS CPS II data (Pope et al. 2002). This model is a Cox Poisson regression model, extended to include random effects for the clustering of geographic units (Ma et al., 2003). Cohort members within a cluster are allowed to have correlated survival times, rather than treating each subject as independent. We specified the random effects as two-level nested clusters, zip codes within counties. Two random effects models were run for each disease outcome and pollutant: (i) a two-level

independent model that treats the effects between units as uncorrelated (i.e., although survival times of cohort members may correlate within a zip code, they are independent of those in any other zip code), and (ii) a two-level dependent model, allowing subjects within a zip code to be correlated with those in adjacent zip codes.

We undertook several additional sensitivity analyses. We limited the PM2.5 analysis to measured values only (1999 through 2002), eliminating estimated PM2.5 data from prior years. This involved a significant reduction in statistical power; however, it is likely that this was counterbalanced by reduced measurement error. Second, we ran two-pollutant models for the major pollutants. Third, we examined two-pollutant models for PM2.5 and ozone, with both limited only to the period 1999 through 2002. Finally, we examined the impact of using values of ozone measured only during the third quarter (summer), when people are likely to spend more time outdoors and to have their windows open.

We performed the Cox proportional hazards analyses using SAS software (SAS Institute, Inc., 2000), and the random effects modeling with the program Cox-Poisson (v. 9.06) provided by Dr. Edward Hughes, invoked through the R programming language (R Development Core Team, 2006).

Results

Table 2 presents descriptive statistics for the members of the study population whose addresses were geocodeable, and includes separate data for movers and nonmovers. At baseline, these participants were predominantly non-Hispanic white (87%), about two-thirds of whom were never-smokers, while five percent were current smokers. Approximately half the population reported second-hand smoke (SHS) exposure at home. A majority of the population reported having a normal or low weight (i.e., reflected in the BMI variable). With few exceptions, movers and nonmovers tended to be quite similar. The exceptions include the following: (1) movers tended to be younger, with 32% under age 40 and 26% age 60 and above, while the corresponding percentages for nonmovers were 11% and 35%; (2) as a consequence, movers were more likely to be pre/perimenopausal (51% versus 35% for nonmovers) and were less likely to have used hormonal therapy (However, this variable had a much greater percentage of unknown/missing responses than any of the other covariates.); (3) movers were slightly less likely to be married or living with a partner; (4) movers reported engaging in slightly more physical activity; and (5) movers were less likely to report having had a family history of AMI or stroke.

Table 3 displays the county of residence for the cohort at baseline. As would be expected, the residential distribution of the study participants reflects that of California's population as a whole, with the majority of the cohort living in the populous counties of Southern California. Table 4 summarizes the exposure data used in the study. For example, aggregating over all of the individual estimates, the long-term (1988 through 2002) average of 24-hour average values of PM2.5 was 22.0 $\mu g/m^3$ with an interquartile range (IQR) of 8.5 $\mu g/m^3$ and an overall range of 3.9 to 38.9 $\mu g/m^3$. For the period 1995 through 2002, the mean PM2.5 level dropped to 18.2 $\mu g/m^3$ with an IQR of 9.3 $\mu g/m^3$, while for the period 1999 through 2002, which included only measured values of PM2.5, the mean was 17.2 $\mu g/m^3$ with an IQR of 9.1 $\mu g/m^3$. The mean one-hour maximum

ozone concentration using data from 1988 through 2002 was 52.1 ppb, with an IQR of 16.4 ppb; the corresponding values for 1995 through 2002 were 48.0 ppb and 12.3). Descriptive statistics are also provided for NO_2, NOx, SO_2, and CO.

Table 2. Percentage distributions of covariates at baseline among members of the CTS cohort whose addresses were geocoded

	Total cohort	Nonmovers	Movers
	N=124,614	N=88,666	N=35,948
	%	%	%
Age (years):			
20-29	4	2	10
30-39	13	9	22
40–49	26	27	23
50–59	24	27	19
60–69	17	19	11
70–79	11	12	9
≥ 80	5	4	6
Race/ethnicity:			
Non-Hispanic White	87	87	86
Other	11	11	12
Unknown/Missing	2	2	2
Smoking:			
Never Smokers	66	65	67
Current smoker	5	5	5
Former smoker	28	29	27
Unknown/Missing	1	1	1
Total smoking pack-years among current and former smokers	33	34	32
BMI (kg/m^2):			
< 18.5	3	3	3
18.5 – 24.9	56	54	58
25 – 29.9	24	25	22
30 – 39	12	12	11
≥ 40	1	2	2
Unknown/Missing	4	4	4
Marital status:			
Married/Living with partner	44	46	40
Divorced/Widowed/Separated/Never Married	21	21	21
Unknown/Missing	35	33	39

Table 2. Percentage distributions of covariates at baseline among members of the CTS cohort whose addresses were geocoded (continued)

Alcohol consumption:			
No alcohol consumption	32	32	32
Beer (yes)	24	23	27
Wine (yes)	56	56	54
Liquor (yes)	30	30	30
Unknown/Missing	6	5	7
SHS adult home exposure:			
No SHS exposure	45	44	48
SHS exposure	49	50	46
Unknown/Missing	6	6	6
Dietary fat (g/day):			
< 41.80	30	31	30
41.80-62.74	30	30	29
≥ 62.75	30	30	31
Unknown/Missing	10	9	10
Dietary fiber g/day):			
< 11.08	30	30	29
11.08-16.01	30	30	30
≥ 16.02	30	30	30
Unknown/Missing	10	10	11
Dietary calories (kcal/day):			
< 1281.63	30	31	28
1281.63-1718.57	30	31	29
≥ 1718.58	30	29	32
Unknown/Missing	10	9	11
Physical activity (hours/week):			
< 2.00	29	30	26
2.00-4.90	34	34	33
≥ 4.91	33	32	37
Unknown/Missing	4	4	4
Menopausal status:			
Pre/peri-menopausal	39	35	51
Post menopausal	52	56	41
Unknown/Missing	9	9	8
Hormone therapy use:			
Never used	17	18	13
Yes	45	48	37
Unknown/Missing	38	34	50

Table 2. Percentage distributions of covariates at baseline among members of the CTS cohort whose addresses were geocoded (continued)

Family history of AMI:			
No	67	65	71
Yes	33	35	29
Family history of stroke:			
No	79	78	82
Yes	21	22	18
Blood pressure medication:			
No regular use	80	79	83
Low	1	1	1
High	15	16	13
Unknown/Missing	4	4	3
Aspirin use:			
No regular use	76	75	78
Low	10	11	10
High	11	11	9
Unknown/Missing	3	3	3

Table 3. California Teachers Study participants' counties of residence at baseline (1995)

COUNTY	COUNT	COUNTY	COUNT
ALAMEDA	5292	ORANGE	11528
ALPINE	7	PLACER	1309
AMADOR	187	PLUMAS	165
BUTTE	1197	RIVERSIDE	5064
CALAVERAS	232	SACRAMENTO	4682
COLUSA	82	SAN BENITO	185
CONTRA COSTA	4480	SAN BERNARDINO	5886
DEL NORTE	80	SAN DIEGO	10876
EL DORADO	868	SAN FRANCISCO	1729
FRESNO	3573	SAN JOAQUIN	2030
GLENN	115	SAN LUIS OBISPO	1491
HUMBOLDT	821	SAN MATEO	2595
IMPERIAL	462	SANTA BARBARA	1728
INYO	109	SANTA CLARA	6289
KERN	2511	SANTA CRUZ	1576
KINGS	409	SHASTA	781
LAKE	229	SIERRA	16
LASSEN	127	SISKIYOU	279
LOS ANGELES	26163	SOLANO	1363
MADERA	472	SONOMA	2327
MARIN	1362	STANISLAUS	1971
MARIPOSA	110	SUTTER	370
MENDOCINO	582	TEHAMA	204
MERCED	831	TRINITY	49
MODOC	42	TULARE	1535
MONO	51	TUOLUMNE	318
MONTEREY	1547	VENTURA	3260
NAPA	763	YOLO	802
NEVADA	650	YUBA	163

Table 4. Descriptive statistics for air pollutants used to estimate long-term exposures among participants in the California Teachers Study

Pollutant	Units	Time period of pollutant average	Mean (SD)	Inter-quartile range	Min-Max Range
Ozone	ppb	1988-2002	52.13 (11.41)	16.41	26.13-88.92
Ozone	ppb	1995-2002	47.99 (9.14)	12.26	20.97-107.93
PM2.5	$\mu g/m^3$	1988-2002	22.00 (5.43)	8.45	3.85-38.89
PM2.5	$\mu g/m^3$	1995-2002	18.15 (5.23)	9.29	3.93-58.86
PM2.5	$\mu g/m^3$	1999-2002	17.24 (4.94)	9.06	4.86-29.76
PM10	$\mu g/m^3$	1988-2002	35.37 (10.95)	14.78	11.39-73.27
PM10	$\mu g/m^3$	1995-2002	31.59 (10.52)	16.56	9.00-81.80
NO_2	ppb	1988-2002	39.85 (12.79)	18.39	6.14-76.76
NO_2	ppb	1995-2002	35.48 (10.93)	17.11	5.54-86.90
NO_X	ppb	1988-2002	104.39 (38.41)	49.10	8.47-240.31
NO_X	ppb	1995-2002	93.58 (35.86)	48.70	7.88-278.98
SO_2	ppb	1988-2002	1.95 (0.74)	0.94	0.25-4.06
SO_2	ppb	1995-2002	1.95 (0.66)	0.86	0.19-4.38
CO	ppm	1988-2002	1.32 (0.46)	0.54	0.41-3.26
CO	ppm	1995-2002	1.05 (0.39)	0.44	0.30-3.48

Table 5 summarizes inter-pollutant correlations for the period 1995 through 2002. These represented the correlations among the estimated exposures for the participants, not the monitored concentrations. For example, PM2.5 was highly correlated with NO_2, NOx and CO (r = 0.79, 0.64, and 0.68, respectively), moderately correlated with ozone (r = 0.41) and least correlated with SO_2 (r = 0.19). Table 6 summarizes the descriptive data and definitions of various traffic metrics used in the analysis.

Table 7 displays estimated hazard ratios for cardiovascular mortality for non-pollutant variables included in the final multivariate models (i.e., for the participants who had PM2.5 data available), both for the entire study population, and disaggregated by residential mobility during the period 1995 through 2002. The HRs for known risk factors for cardiovascular mortality are generally in the expected directions, e.g., current smoking, BMI, marital status, alcohol consumption, SHS exposure, dietary fat, fiber, and calories, physical activity, and menopausal status. The decreased risk associated with hormone therapy is consistent with some observational studies, but not with several controlled clinical trials (Alexandersen et al. 2006; Hedblad et al. 2002; Waters et al. 2002; Garbe et al. 2004). Interestingly, family history of stroke or AMI was not associated with cardiovascular mortality in this analysis. A number of these variables

appeared more strongly associated with cardiovascular mortality among the movers than nonmovers, especially current smoking, extreme obesity (BMI > 40), single marital status, and low dietary fiber. Table 7 also includes HRs for a variety of contextual variables in the regression models; none was associated with cardiovascular mortality except the Census block group percent unemployed and percentage with at least a bachelor's degree: both of these associations were significant only among movers.

Table 5. Spearman correlation coefficients (r) for estimated pollutant exposures among CTS participants for the period 1995-2002

	Ozone	PM2.5 r p-value n	PM10 r p-value n	NO_2 r p-value n	CO r p-value n	NOx r p-value n	SO_2 r p-value n
Ozone r p-value n	1.00000 106,409	0.4134 <.0001 97,283	0.62304 <.0001 68,238	0.38842 <.0001 16,636	0.10939 <.0001 10,632	0.05219 <.0001 16,497	-0.06386 <.0001 20,785
PM2.5 r p-value n		1.00000 98,426	0.86369 <.0001 64,316	0.79392 <.0001 16,064	0.67839 <.0001 10,057	0.64070 <.0001 15,935	0.19461 <.0001 20,529
PM10 r p-value n			1.00000 68,957	0.76376 <.0001 14,229	0.51587 <.0001 8,904	0.50481 <.0001 14,090	0.01761 0.0165 15,995
NO_2 r p-value n				1.00000 16,636	0.78095 <.0001 8,900	0.86786 <.0001 16,497	0.00175 <.0001 8,880
CO r p-value n					1.00000 10,632	0.81258 <.0001 8,761	0.07537 <.0001 5,713
NOx r p-value n						1.00000 16,497	0.03659 0.0505 8,880
SO_2 r p-value n							1.00000 20,785

Table 6. Descriptive statistics for traffic, vehicle and road measures, California Teachers Study Cohort

Variable*	N	25th percentile	75th percentile	IQR
Distance to highway	109,039	657.728	2479.08E	1821.352
Traffic density 150m	109,039	0.10442	1430.03499	1429.93057
Traffic density 300m	109,039	103.61988	1969.4252	1865.80532
Vehicle density	109,039	774	1602	828
Road density	109,018	698.87063	1121.29834	422.42771

* Traffic variable definitions:

Distance to highway= Proximity of residence to a "major" highway, in meters. (Limited to within 20km.) Missing data (n=180) changed to 49999.

Traffic density 150m = Vehicle Miles Traveled within 150 meters of a residence using conflated TeleAtlas 2005q2 centerlines linked to HPMS 2000. Missing values (n=46,909) set to minimum non-zero value (0.10442).

Traffic density 300m = Vehicle Miles Traveled within 300 meters of a residence using conflated TeleAtlas 2005q2 centerlines linked to HPMS 2000. (Normalized to 150m values.) Missing values (n=19,700) set to minimum non-zero value (0.00339).

Road Density = Meters of roads (based on TeleAtlas/Dynamap road data) within 150 meters of a residence

Vehicle Density = 2000 Census Block group count of aggregate number of vehicles available from occupied housing units.

Table 7. Hazard ratios for nonpollutant covariates in relation to cardiovascular mortality for CTS participants with PM2.5 data available

	Total cohort		Non-movers		Movers	
	N=98,426, # deaths = 2,296		N=72,152; # deaths = 1,386		N=26,274, # deaths = 910	
	%	HR (95% CI)	%	HR (95% CI)	%	HR (95% CI)
Smoking status						
Never smokers	67	1.000	67	1.000	68	1.000
Current smoker	5	1.239 (1.081, 1.419)	5	1.176 (0.990, 1.396)	5	1.466 (1.169, 1.837)
Former smoker	28	1.006 (0.928, 1.091)	28	0.987 (0.891, 1.094)	27	1.051 (0.918, 1.202)
Total smoking pack-years among current and former smokers	33	1.010 (1.008, 1.012)	33	1.011 (1.009, 1.014)	32	1.007 (1.004, 1.010)
BMI (kg/m^2)						
< 18.5	3	1.732 (1.509, 1.988)	3	1.757 (1.474, 2.094)	3	1.661 (1.323, 2.086)
18.5 – 24.9	55	1.093 (1.016, 1.177)	54	1.056 (0.964, 1.158)	58	1.120 (0.988, 1.270)
25 – 29.9	24	1.000	25	1.000	22	1.000
30 – 39	12	1.027 (0.920, 1.146)	12	1.025 (0.895, 1.173)	11	1.013 (0.837, 1.226)
≥ 40	2	1.772 (1.397, 2.248)	2	1.459 (1.068, 1.994)	2	2.569 (1.768, 3.732)
Unknown/Missing	4	1.014 (0.912, 1.128)	4	1.026 (0.895, 1.175)	4	0.934 (0.785, 1.111)
Marital status						
Married/Living with partner	44	1.000	46	1.000	39	1.000
Divorced/Widowed/Separated/Never Married	22	1.244 (1.094, 1.414)	21	1.098 (0.942, 1.279)	23	1.608 (1.255, 2.061)
Unknown/Missing	34	7.095 (6.384, 7.885)	33	6.480 (5.739, 7.317)	38	8.308 (6.674, 10.342)
Alcohol consumption						
No alcohol consumption	33	1.000	33	1.000	33	1.000
Beer (yes)	24	0.940 (0.857, 1.030)	23	0.928 (0.828, 1.039)	26	0.969 (0.828, 1.133)
Wine (yes)	57	0.862 (0.806, 0.923)	57	0.839 (0.770, 0.915)	56	0.898 (0.804, 1.004)
Liquor (yes)	30	0.885 (0.822, 0.952)	30	0.917 (0.836, 1.007)	30	0.865 (0.764, 0.979)
Unknown/Missing	4	0.943 (0.828, 1.076)	4	0.944 (0.793, 1.124)	4	0.958 (0.782, 1.172)

Table 7. Hazard ratios for nonpollutant covariates in relation to cardiovascular mortality for CTS participants with PM2.5 data available (continued)

	Total cohort		Non-movers		Movers	
	%	HR (95% CI)	%	HR (95% CI)	%	HR (95% CI)
SHS exposure at home						
No SHS exposure	45	1.000	45	1.000	47	1.000
SHS exposure	50	1.072 (1.002, 1.148)	50	1.038 (0.953, 1.131)	49	1.097 (0.979, 1.230)
Unknown/Missing	5	1.023 (0.890, 1.176)	5	1.006 (0.842, 1.201)	4	1.036 (0.824, 1.303)
Dietary fat (g/day)						
< 41.80	31	1.000	31	1.000	30	1.000
41.80-62.74	31	1.084 (0.994, 1.183)	31	1.073 (0.962, 1.195)	30	1.131 (0.974, 1.314)
≥ 62.75	30	1.214 (1.079, 1.367)	30	1.204 (1.039, 1.395)	31	1.256 (1.024, 1.542)
Unknown/Missing	8	1.286 (1.127, 1.466)	8	1.224 (1.034, 1.450)	9	1.401 (1.132, 1.735)
Dietary fiber (g/day)						
≥ 16.02	30	1.000	30	1.000	31	1.000
11.08-16.01	31	1.024 (0.938, 1.118)	31	0.985 (0.883, 1.099)	30	1.080 (0.933, 1.250)
< 11.08	31	1.182 (1.074, 1.300)	31	1.102 (0.978, 1.243)	30	1.386 (1.181, 1.627)
Unknown/Missing	8	1.286 (1.127, 1.466)	8	1.224 (1.034, 1.450)	9	1.401 (1.132, 1.735)
Dietary calories (kcal/day)						
< 1281.63	31	1.000	31	1.000	30	1.000
1281.63 – 1718.57	31	1.014 (0.921, 1.117)	31	1.026 (0.911, 1.157)	30	1.011 (0.857, 1.193)
≥ 1718.58	30	1.072 (0.933, 1.232)	30	1.045 (0.878, 1.243)	31	1.114 (0.880, 1.410)
Unknown/Missing dietary information	8	1.286 (1.127, 1.466)	8	1.224 (1.034, 1.450)	9	1.401 (1.132, 1.735)
Physical activity (hr/wk)						
≥ 4.91	32	1.000	31	1.000	36	1.000
2.00-4.90	34	1.026 (0.946, 1.114)	34	1.038 (0.936, 1.150)	33	1.012 (0.883, 1.159)
< 2.00	30	1.061 (0.983, 1.146)	31	1.099 (0.997, 1.211)	27	0.986 (0.868, 1.120)
Unknown/Missing	4	0.993 (0.892, 1.105)	4	1.065 (0.930, 1.220)	4	0.902 (0.756, 1.076)

Table 7. Hazard ratios for nonpollutant covariates in relation to cardiovascular mortality for CTS participants with PM2.5 data available (continued)

	Total cohort		Non-movers		Movers	
	%	HR (95% CI)	%	HR (95% CI)	%	HR (95% CI)
Menopausal status						
Pre/peri-menopausal	37	1.000	34	1.000	47	1.000
Post menopausal	54	1.412 (1.080, 1.846)	57	1.330 (0.968, 1.829)	55	1.340 (0.796, 2.255)
Unknown/Missing	9	0.953 (0.702, 1.293)	9	0.853 (0.591, 1.231)	8	1.039 (0.590, 1.831)
Hormone therapy use						
Never used	17	1.000	19	1.000	14	1.000
Yes	47	0.866 (0.814, 0.922)	49	0.836 (0.773, 0.905)	40	0.903 (0.814, 1.002)
Unknown/Missing	36	0.821 (0.689, 0.977)	32	0.703 (0.555, 0.889)	46	0.959 (0.738, 1.248)
Family history of AMI						
No	66	1.000	64	1.000	69	1.000
Yes	34	0.990 (0.931, 1.052)	36	1.026 (0.950, 1.108)	31	0.949 (0.857, 1.052)
Stroke Family history of stroke						
No	78	1.000	77	1.000	81	1.000
Yes	22	1.014 (0.949, 1.083)	23	1.079 (0.993, 1.172)	19	0.895 (0.801, 1.000)
Blood pressure medication:						
No regular use	80	1.000	79	1.000	82	1.000
Low	1	1.809 (1.468, 2.230)	1	2.053 (1.608, 2.622)	1	1.544 (1.024, 2.329)
High	16	1.385 (1.297, 1.479)	17	1.358 (1.249, 1.477)	14	1.420 (1.276, 1.580)
Unknown/Missing	3	1.325 (1.183, 1.485)	3	1.406 (1.212, 1.632)	3	1.222 (1.021, 1.463)
Aspirin use						
No regular use	76	1.000	75	1.000	77	1.000
Low	10	0.928 (0.829, 1.040)	11	0.953 (0.829, 1.095)	10	0.925 (0.760, 1.126)
High	11	1.088 (1.011, 1.172)	11	1.124 (1.024, 1.234)	10	1.035 (0.915, 1.171)
Unknown/Missing	3	0.924 (0.801, 1.065)	3	0.886 (0.733, 1.071)	3	0.958 (0.771, 1.191)

Table 7. Hazard ratios for nonpollutant covariates in relation to cardiovascular mortality for CTS participants with PM2.5 data available (continued)

	Total cohort	Non-movers	Movers		Total cohort	Non-movers
	%	HR (95% CI)	%		%	HR (95% CI)
Contextual effects (1990 census variables at the block group level)						
Household median income	-	1.000 (1.000, 1.000)	-	1.000 (1.000, 1.000)	-	1.000 (1.000, 1.000)
Percent in poverty	-	0.994 (0.988, 0.999)	-	0.994 (0.987, 1.001)	-	0.993 (0.984, 1.002)
Percent with at least a bachelor's degree	-	1.000 (0.998, 1.003)	-	0.998 (0.995, 1.001)	-	1.006 (1.002, 1.010)
Total population	-	1.000 (1.000, 1.000)	-	1.000 (1.000, 1.000)	-	1.000 (1.000, 1.000)
Percent black	-	0.998 (0.994, 1.003)	-	0.998 (0.992, 1.003)	-	1.001 (0.994, 1.009)
Percent non-Hispanic white	-	1.000 (0.997, 1.003)	-	1.000 (0.996, 1.004)	-	0.999 (0.993, 1.004)
Percent Hispanic	-	1.000 (0.996, 1.005)	-	0.999 (0.993, 1.005)	-	1.002 (0.995, 1.009)
Percent unemployed	-	1.007 (0.999, 1.016)	-	1.004 (0.993, 1.015)	-	1.012 (1.000, 1.025)

Tables 8a and 8b summarize the results for associations of all-cause and cardiovascular mortality, as well as incidence of AMI and stroke, with PM2.5, using exposures estimated from 1995 through 2002 and 1988 through 2002, respectively. These results are presented in terms of the increase in the HR per unit increase in $\mu g/m^3$, and are provided for the full cohort and separately for the movers and nonmovers. In addition, the tables indicate the results for the model including the pollution term (adjusted only for age and race/ethnicity) and the full models that include the individual-level and contextual (Census block-based) covariates. Additional covariates were included in the regression models for cardiopulmonary mortality and incidence of AMI and stroke, as described in the methods section (i.e., family history of MI or stroke, use of medication for hypertension, and long-term use of aspirin). The results reveal the following: (1) strong and highly significant associations between PM2.5 and all four outcomes; (2) little quantitative difference in results for the model including only the pollutant term plus age and race versus the full models including both individual-level and contextual covariates; (3) in most cases, little difference in risk between the full cohort and the movers or nonmovers; and (4) little difference between the shorter (1995-2002) and longer (1988-2002) exposure periods.

Tables 9a and 9b summarize the results for PM10. For this pollutant, the results suggest the following: (1) associations between PM10 and all four adverse health outcomes (all-cause and cardiopulmonary mortality, incidence of AMI and stroke) are evident for the 1988 to 2002 exposure period, but not when the exposure period is limited to 1995 to 2002; (2) there was little quantitative difference between the results for the simple model (pollutant + age + race) versus the full model including both individual-level and contextual covariates; and (3) there was little difference, in most cases, between the full cohort and the sub-cohorts of movers or nonmovers. The only significant quantitative difference occurred for stroke incidence, for which associations are evident (using 1988-2002 exposures) for nonmovers but not for movers (Table 9b).

Tables 10a and 10b summarize the results for ozone. As with PM2.5 and PM10, there was little quantitative difference between the results obtained with the simple versus the full model, though for AMI incidence in both exposure periods, the ozone coefficients decreased slightly in the full model. In addition, during the period 1995 through 2002, the coefficients linking ozone with cardiopulmonary mortality went from significant to nonsignificant when the full model instead of the simple model was used, and the coefficient for stroke incidence among nonmovers declined from borderline significant ($p = 0.08$) to nonsignificant ($p = 0.31$). In general, the associations between ozone and the various outcomes appeared to be slightly stronger among nonmovers than movers, except for all-cause mortality from 1995 through 2002. While there were strong associations between ozone and all four outcomes during the longer exposure period (1988-2002), only AMI incidence was associated with ozone during the period 1995 through 2002 in both the simple and full models. As was observed in the single-pollutant model for PM10, during the exposure period 1988 through 2002, ozone was strongly associated with all four outcomes (except for stroke incidence among movers).

Table 8a. Results for PM2.5 using exposures from 1995-2002 for the full cohort, nonmovers and movers, disaggregated by model specification

Model	All cohort Baseline address	Nonmovers	Movers
All-cause mortality			
N	98,426	72,152	26,274
# deaths	4,783	2,980	1,803
Mean PM2.5 (SD) (ug/m^3)	18.15 (5.23)	18.12 (5.24)	18.14 (5.12)
Mean follow-up time (years)	6.70	6.72	6.63
	HR 95% CI p-value	HR 95% CI p-value	HR 95% CI p-value
Age/race strata	1.048 (1.042, 1.053) <0.0001	1.049 (1.042, 1.057) <0.0001	1.056 (1.046, 1.065) <0.0001
Age/race strata + covariates*	1.051 (1.045, 1.057) <0.0001	1.053 (1.046, 1.061) <0.0001	1.053 (1.043, 1.062) <0.0001
Cardiopulmonary mortality			
N	98,426	72,152	26,274
# deaths	2,296	1,386	910
Mean PM2.5 (SD) (ug/m^3)	18.15 (5.23)	18.12 (5.24)	18.14 (5.12)
Mean follow-up time (years)	6.70	6.72	6.63
	HR 95% CI p-value	HR 95% CI p-value	HR 95% CI p-value
Age/race strata	1.048 (1.040, 1.056) <0.0001	1.049 (1.039, 1.060) <0.0001	1.057 (1.044, 1.071) <0.0001
Age/race strata + covariates**	1.050 (1.042, 1.059) <0.0001	1.052 (1.041, 1.063) <0.0001	1.054 (1.040, 1.068) <0.0001
AMI incidence			
N	97,750	71,667	26,083
# events	1,966	1,332	634
Mean PM2.5 (SD) (ug/m^3)	18.15 (5.23)	18.12 (5.23)	18.14 (5.12)
Mean follow-up time (years)	6.70	6.73	6.63
	HR 95% CI p-value	HR 95% CI p-value	HR 95% CI p-value
Age/race strata	1.056 (1.047, 1.065) <0.0001	1.054 (1.044, 1.065) <0.0001	1.069 (1.053, 1.085) <0.0001
Age/race strata + covariates**	1.059 (1.050, 1.068) <0.0001	1.056 (1.044, 1.067) <0.0001	1.070 (1.054, 1.087) <0.0001
Stroke incidence			
N	98,017	71,871	26,146
# events	1,379	935	444
Mean PM2.5 (SD) (ug/m^3)	18.15 (5.23)	18.12 (5.23)	18.14 (5.12)
Mean follow-up time (years)	6.70	6.73	6.63
	HR 95% CI p-value	HR 95% CI p-value	HR 95% CI p-value
Age/race strata	1.058 (1.047, 1.069) <0.0001	1.052 (1.039, 1.065) <0.0001	1.080 (1.061, 1.099) <0.0001
Age/race strata + covariates**	1.061 (1.050, 1.072) <0.0001	1.054 (1.040, 1.068) <0.0001	1.081 (1.061, 1.101) <0.0001

* Adjusted for smoking status, total pack years, BMI, marital status, alcohol consumption, SHS home exposure, dietary fat, dietary fiber, dietary calories, physical activity, menopausal status, HT use; and contextual variables (income, income inequality, education, population size, racial composition, unemployment).

** Includes the above risk factors plus: family history of AMI, family history of stroke, blood pressure medication and aspirin use.

Table 8b. Results for PM2.5 using exposures from 1988-2002 for the full cohort, nonmovers and movers, disaggregated by model specification

Model	All cohort Baseline address	Nonmovers	Movers
All-cause (natural) mortality			
N	98,426	72,152	26,274
# deaths	4,783	2,980	1,803
Mean PM2.5 (SD) (ug/m^3)	22.00 (5.43)	21.96 (5.42)	22.18 (5.32)
Mean follow-up time (years)	6.70	6.72	6.63
	HR 95% CI p-value	HR 95% CI p-value	HR 95% CI p-value
Age/race strata	1.051 (1.045, 1.056) <0.0001	1.053 (1.046, 1.060) <0.0001	1.053 (1.043, 1.062) <0.0001
Age/race strata + covariates*	1.052 (1.046, 1.058) <0.0001	1.055 (1.048, 1.062) <0.0001	1.049 (1.039, 1.059) <0.0001
Cardiopulmonary mortality			
N	98,426	72,152	26,274
# deaths	2,296	1,386	910
Mean PM2.5 (SD) (ug/m^3)	22.00 (5.43)	21.96 (5.42)	22.18 (5.32)
Mean follow-up time (years)	6.70	6.72	6.63
	HR 95% CI p-value	HR 95% CI p-value	HR 95% CI p-value
Age/race strata	1.050 (1.042, 1.058) <0.0001	1.053 (1.043, 1.063) <0.0001	1.052 (1.039, 1.065) <0.0001
Age/race strata + covariates**	1.051 (1.043, 1.059) <0.0001	1.054 (1.043, 1.065) <0.0001	1.048 (1.034, 1.061) <0.0001
AMI incidence			
N	97,750	71,667	26,083
# events	1,966	1,332	634
Mean PM2.5 (SD) (ug/m^3)	21.99 (5.43)	21.95 (5.42)	22.19 (5.32)
Mean follow-up time (years)	6.70	6.73	6.63
	HR 95% CI p-value	HR 95% CI p-value	HR 95% CI p-value
Age/race strata	1.057 (1.048, 1.065) <0.0001	1.058 (1.047, 1.069) <0.0001	1.059 (1.043, 1.075) <0.0001
Age/race strata + covariates**	1.059 (1.050, 1.068) <0.0001	1.059 (1.048, 1.071) <0.0001	1.060 (1.043, 1.076) <0.0001
Stroke incidence			
N	98,017	71,871	26,146
# events	1,379	935	444
Mean PM2.5 (SD) (ug/m^3)	22.00 (5.43)	21.96 (5.42)	22.19 (5.32)
Mean follow-up time (years)	6.70	6.73	6.63
	HR 95% CI p-value	HR 95% CI p-value	HR 95% CI p-value
Age/race strata	1.053 (1.043, 1.064) <0.0001	1.051 (1.038, 1.063) <0.0001	1.065 (1.046, 1.084) <0.0001
Age/race strata + covariates**	1.055 (1.044, 1.066) <0.0001	1.052 (1.039, 1.065) <0.0001	1.064 (1.044, 1.084) <0.0001

* Adjusted for smoking status, total pack years, BMI, marital status, alcohol consumption, SHS home exposure, dietary fat, dietary fiber, dietary calories, physical activity, menopausal status, HT use; and contextual variables (income, income inequality, education, population size, racial composition, unemployment).

** Includes the above risk factors plus: family history of AMI, family history of stroke, blood pressure medication and aspirin use.

Table 9a. Results for PM10 using exposures from 1995-2002 for the full cohort, nonmovers and movers, disaggregated by model specification

Model	All cohort Baseline address	Nonmovers	Movers
All-cause (natural) mortality			
N	68,957	50,256	18,701
# deaths	3,525	2,131	1,394
Mean PM10 (sd) ($\mu g/m^3$)	31.59 (10.52)	31.41 (10.45)	31.79 (10.54)
Mean follow-up time (years)	6.68	6.71	6.61
	HR 95% CI p-value	HR 95% CI p-value	HR 95% CI p-value
Age/race strata	0.999 (0.996, 1.003) 0.730	1.000 (0.996, 1.005) 0.843	1.000 (0.995, 1.005) 0.959
Age/race strata + covariates*	0.999 (0.995, 1.002) 0.458	1.000 (0.996, 1.004) 0.965	1.000 (0.994, 1.005) 0.936
Cardiopulmonary mortality			
N	68,957	50,256	18,701
# deaths	1,739	1,025	714
Mean PM10 ($\mu g/m^3$)	31.59 (10.52)	31.41 (10.45)	31.79 (10.54)
Mean follow-up time (years)	6.68	6.71	6.61
	HR 95% CI p-value	HR 95% CI p-value	HR 95% CI p-value
Age/race strata	1.001 (0.996, 1.005) 0.754	1.002 (0.996, 1.008) 0.607	1.002 (0.995, 1.009) 0.614
Age/race strata + covariates**	1.000 (0.995, 1.005) 0.976	1.002 (0.995, 1.008) 0.589	1.002 (0.994, 1.010) 0.675
AMI incidence			
N	68,477	49,907	18,570
# events	1,460	953	507
Mean PM10 (µg/m3)	31.59 (10.52)	31.41 (10.45)	31.81 (10.55)
Mean follow-up time (years)	6.69	6.72	6.61
	HR 95% CI p-value	HR 95% CI p-value	HR 95% CI p-value
Age/race strata	1.003 (0.998, 1.008) 0.208	1.001 (0.995, 1.008) 0.658	1.008 (0.999, 1.017) 0.066
Age/race strata + covariates**	1.002 (0.997, 1.007) 0.415	1.001 (0.994, 1.007) 0.797	1.006 (0.997, 1.016) 0.165
Stroke incidence			
N	68,671	50,055	18,616
# events	1,040	684	356
Mean PM10 (µg/m3)	31.60 (10.52)	31.42 (10.45)	31.80 (10.55)
Mean follow-up time (years)	6.69	6.72	6.61
	HR 95% CI p-value	HR 95% CI p-value	HR 95% CI p-value
Age/race strata	1.003 (0.998, 1.009) 0.258	1.005 (0.998, 1.012) 0.160	1.002 (0.992, 1.013) 0.641
Age/race strata + covariates**	1.002 (0.996, 1.008) 0.582	1.004 (0.996, 1.011) 0.355	1.001 (0.991, 1.012) 0.804

* Adjusted for smoking status, total pack years, BMI, marital status, alcohol consumption, SHS home exposure, dietary fat, dietary fiber, dietary calories, physical activity, menopausal status, HT use; and contextual variables (income, income inequality, education, population size, racial composition, unemployment).

** Includes the above risk factors plus: family history of AMI, family history of stroke, blood pressure medication and aspirin use.

Table 9b. Results for PM10 using exposures from 1988-2002 for the full cohort, nonmovers and movers, disaggregated by model specification

Model	All cohort Baseline address	Non-movers	Movers
All-cause (natural) mortality			
N	68,957	50,256	18,701
# deaths	3,525	2,131	1,394
Mean PM10 ($\mu g/m^3$)	35.37 (10.95)	35.15 (10.85)	36.18 (11.10)
Mean follow-up time (years)	6.68	6.71	6.61
	HR 95% CI p-value	HR 95% CI p-value	HR 95% CI p-value
Age/race strata	1.008 (1.005, 1.011) <0.0001	1.009 (1.005, 1.013) <0.0001	1.006 (1.001, 1.011) 0.017
Age/race strata + covariates*	1.008 (1.004, 1.011) <0.0001	1.009 (1.005, 1.013) <0.0001	1.006 (1.001, 1.011) 0.027
Cardiopulmonary mortality			
N	68,957	50,256	18,701
# deaths	1,739	1,025	714
Mean PM10 ($\mu g/m^3$)	35.37 (10.95)	35.15 (10.85)	36.18 (11.10)
Mean follow-up time (years)	6.68	6.71	6.61
	HR 95% CI p-value	HR 95% CI p-value	HR 95% CI p-value
Age/race strata	1.009 (1.005, 1.014) <0.0001	1.010 (1.004, 1.016) 0.0005	1.008 (1.001, 1.015) 0.021
Age/race strata + covariates**	1.009 (1.004, 1.013) 0.0002	1.010 (1.004, 1.016) 0.001	1.008 (1.000, 1.015) 0.042
AMI incidence			
N	68,477	49,907	18,570
# events	1,460	953	507
Mean PM10 ($\mu g/m^3$)	35.37 (10.94)	35.15 (10.84)	36.19 (11.10)
Mean follow-up time (years)	6.69	6.72	6.61
	HR 95% CI p-value	HR 95% CI p-value	HR 95% CI p-value
Age/race strata	1.013 (1.008, 1.007) <0.0001	1.012 (1.006, 1.018) <0.0001	1.014 (1.006, 1.022) 0.0006
Age/race strata + covariates**	1.012 (1.007, 1.007) <0.0001	1.012 (1.006, 1.018) 0.0001	1.013 (1.004, 1.021) 0.003
Stroke incidence			
N	68,671	50,055	18,616
# events	1,040	684	356
Mean PM10 ($\mu g/m^3$)	35.38 (10.95)	35.16 (10.85)	36.18 (11.11)
Mean follow-up time (years)	6.69	6.72	6.61
	HR 95% CI p-value	HR 95% CI p-value	HR 95% CI p-value
Age/race strata	1.011 (1.005, 1.016) 0.0002	1.012 (1.006, 1.019) 0.0003	1.007 (0.997, 1.017) 0.157
Age/race strata + covariates**	1.009 (1.003, 1.015) 0.002	1.011 (1.004, 1.019) 0.002	1.006 (0.996, 1.016) 0.254

* Adjusted for smoking status, total pack years, BMI, marital status, alcohol consumption, SHS home exposure, dietary fat, dietary fiber, dietary calories, physical activity, menopausal status, HT use; and contextual variables (income, income inequality, education, population size, racial composition, unemployment).

** Includes the above risk factors plus: family history of AMI, family history of stroke, blood pressure medication and aspirin use.

Table 10a. Results for ozone using exposures from 1995-2002 for the full cohort, nonmovers and movers, disaggregated by model specification

Model	All cohort Baseline address			Nonmovers			Movers		
All-cause (natural) mortality									
N	106,409			77,942			28,467		
# deaths	5,219			3,229			1,990		
Mean ozone (ppb)	47.99 (9.14)			47.77 (9.11)			48.80 (8.83)		
Mean follow-up time (years)	6.69			6.72			6.62		
	HR	95% CI	p-value	HR	95% CI	p-value	HR	95% CI	p-value
Age/race strata	1.001	(0.998, 1.004)	0.461	1.001	(0.997, 1.005)	0.576	0.998	(0.993, 1.003)	0.479
Age/race strata + covariates*	1.000	(0.997, 1.003)	0.949	1.000	(0.996, 1.004)	0.962	0.999	(0.994, 1.005)	0.773
Cardiopulmonary mortality									
N	106,409			77,942			28,467		
# deaths	2,516			1,504			1,012		
Mean ozone (ppb)	47.99 (9.14)			47.77 (9.11)			48.80 (8.83)		
Mean follow-up time (years)	6.69			6.72			6.62		
	HR	95% CI	p-value	HR	95% CI	p-value	HR	95% CI	p-value
Age/race strata	1.005	(1.001, 1.010)	0.019	1.006	(1.000, 1.011)	0.055	1.003	(0.996, 1.010)	0.416
Age/race strata + covariates**	1.004	(0.999, 1.009)	0.108	1.005	(0.999, 1.011)	0.104	1.004	(0.997, 1.012)	0.287
AMI incidence									
N	105,695			77,434			28,261		
# events	2,189			1,479			710		
Mean ozone (ppb)	47.99 (9.14)			47.77 (9.11)			48.81 (8.83)		
Mean follow-up time (years)	6.70			6.72			6.63		
	HR	95% CI	p-value	HR	95% CI	p-value	HR	95% CI	p-value
Age/race strata	1.011	(1.007, 1.016)	<0.0001	1.011	(1.005, 1.017)	0.0002	1.010	(1.001, 1.018)	0.023
Age/race strata + covariates**	1.007	(1.002, 1.012)	0.007	1.008	(1.001, 1.014)	0.016	1.007	(0.998, 1.016)	0.139
Stroke incidence									
N	105,981			77,647			28,334		
# events	1,507			1,020			487		
Mean ozone (ppb)	48.00 (9.14)			47.77 (9.11)			48.80 (8.83)		
Mean follow-up time (years)	6.70			6.72			6.63		
	HR	95% CI	p-value	HR	95% CI	p-value	HR	95% CI	p-value
Age/race strata	1.002	(0.997, 1.008)	0.382	1.006	(0.999, 1.013)	0.081	0.993	(0.982, 1.003)	0.153
Age/race strata + covariates**	1.001	(0.995, 1.007)	0.840	1.004	(0.996, 1.011)	0.310	0.993	(0.982, 1.004)	0.240

* Adjusted for smoking status, total pack years, BMI, marital status, alcohol consumption, SHS home exposure, dietary fat, dietary fiber, dietary calories, physical activity, menopausal status, HT use; and contextual variables (income, income inequality, education, population size, racial composition, unemployment).

** Includes the above risk factors plus: family history of AMI, family history of stroke, blood pressure medication and aspirin use.

Table 10b. Results for ozone using exposures from 1988-2002 for the full cohort, nonmovers and movers, disaggregated by model specification

Model	All cohort Baseline address	Nonmovers	Movers
All-cause (natural) mortality			
N	106,409	77,942	28,467
# deaths	5,219	3,229	1,990
Mean ozone (ppb)	52.13 (11.41)	51.89 (11.40)	52.96 (11.22)
Mean follow-up time (years)	6.69	6.72	6.62
	HR 95% CI p-value	HR 95% CI p-value	HR 95% CI p-value
Age/race strata	1.009 (1.006, 1.011) <0.0001	1.009 (1.006, 1.012) <0.0001	1.006 (1.002, 1.010) 0.002
Age/race strata + covariates*	1.009 (1.006, 1.011) <0.0001	1.010 (1.006, 1.013) <0.0001	1.007 (1.003, 1.011) 0.001
Cardiopulmonary mortality			
N	106,409	77,942	28,467
# deaths	2,516	1,504	1,012
Mean ozone (ppb)	52.13 (11.41)	51.89 (11.40)	52.96 (11.22)
Mean follow-up time (years)	6.69	6.72	6.62
	HR 95% CI p-value	HR 95% CI p-value	HR 95% CI p-value
Age/race strata	1.011 (1.008, 1.015) <0.0001	1.013 (1.008, 1.017) <0.0001	1.009 (1.003, 1.015) 0.002
Age/race strata + covariates**	1.011 (1.007, 1.015) <0.0001	1.013 (1.008, 1.017) <0.0001	1.009 (1.003, 1.015) 0.003
AMI incidence			
N	105,695	77,434	28,261
# events	2,189	1,479	710
Mean ozone (ppb)	52.13 (11.41)	51.88 (11.40)	52.97 (11.22)
Mean follow-up time (years)	6.70	6.72	6.63
	HR 95% CI p-value	HR 95% CI p-value	HR 95% CI p-value
Age/race strata	1.017 (1.014, 1.021) <0.0001	1.018 (1.013, 1.022) <0.0001	1.016 (1.009, 1.023) <0.0001
Age/race strata + covariates**	1.015 (1.011, 1.009) <0.0001	1.016 (1.012, 1.021) <0.0001	1.014 (1.007, 1.022) <0.0001
Stroke incidence			
N	105,981	77,647	28,334
# events	1,507	1,020	487
Mean ozone (ppb)	52.13 (11.41)	51.89 (11.40)	52.96 (11.22)
Mean follow-up time (years)	6.70	6.72	6.63
	HR 95% CI p-value	HR 95% CI p-value	HR 95% CI p-value
Age/race strata	1.011 (1.007, 1.016) <0.0001	1.013 (1.007, 1.018) <0.0001	1.007 (0.999, 1.015) 0.111
Age/race strata + covariates**	1.011 (1.006, 1.015) <0.0001	1.012 (1.006, 1.018) <0.0001	1.007 (0.998, 1.015) 0.113

* Adjusted for smoking status, total pack years, BMI, marital status, alcohol consumption, SHS home exposure, dietary fat, dietary fiber, dietary calories, physical activity, menopausal status, HT use; and contextual variables (income, income inequality, education, population size, racial composition, unemployment).

** Includes the above risk factors plus: family history of AMI, family history of stroke, blood pressure medication and aspirin use.

Tables 11a-11c summarize the results for all seven pollutants for four outcomes, using the IQR of exposure for 1995 through 2002 for the full cohort (Table 11a), nonmovers (Table 11b) and movers (Table 11c). For PM2.5, PM10, and ozone, the results are simply restatements from Tables 8 through 10, but expressed in terms of the IQR instead of per unit (either $\mu g/m^3$ or ppb). The tables indicate that there are far fewer events included in the analysis of NO_2, NOx, SO_2, and CO, since the spatial catchment areas for these pollutants were much smaller than for ozone and PM. In addition, there were substantially fewer monitors in the database for these gaseous pollutants, especially SO_2. For the full cohort, NO_2 and CO were associated with all-cause (but not cardiopulmonary) mortality and with the incidence of both AMI and stroke. Neither SO_2 nor NOx was significantly associated with these outcomes in the full cohort. Fewer associations were observed for nonmovers (e.g., NO_2 was associated only with stroke, and CO with all-cause mortality and AMI incidence); however, in this subcohort SO_2 was associated with both all-cause and cardiopulmonary mortality. Among movers, NO_2, NOx and CO were associated with AMI incidence, while CO was also associated with stroke incidence.

Tables 12a-12c summarize a similar set of results using the exposure period 1988 through 2002. The results for PM2.5, PM10, and ozone were discussed above in relation to Tables 8 through 10. In the full cohort, generally similar results emerge for the other gaseous pollutants. NO_2 and CO were associated with all-cause (but not cardiopulmonary) mortality and the incidence of both AMI and stroke, while NOx was associated only with AMI incidence. While the coefficients for ozone and PM10 became stronger and more significant for the longer exposure period, this was not the case for either NO_2 or CO, the coefficients for which were quite similar to those for the 1995-2002 exposure period. Among nonmovers, similar associations were observed for NO_2 (stroke incidence) and CO (all-cause mortality and AMI incidence) as for the shorter exposure period. Among movers, both NO_2 and NOx were both associated with AMI incidence. There were no associations observed for SO_2 in the full cohort or either of the subcohorts for the 1988-2002 exposure period.

Table 13 summarizes the results for traffic, road and vehicle metrics, with hazard ratios calculated per IQR. Neither of the traffic density measures showed a relationship with the outcomes of interest. However, the distance to a major highway was inversely associated with AMI risk. Road density (i.e., meters of roads within 150 meters of the subjects' residences) was associated with both all-cause mortality and stroke incidence. In addition, an association of marginal statistical significance was observed between vehicle density (i.e., 2000 Census block group count of aggregate number of vehicles available from occupied housing units) and stroke incidence.

Table 11a. Hazard ratios for average air pollutant exposures (1995-2002) in relationship to disease outcomes for full CTS cohort based on pollutant interquartile ranges

Pollutant	Units	Outcome	# events	N	IQR	HR (95% CI)
Ozone	ppb	All cause mortality	5,219	106,409	12.26	1.00 (0.96, 1.04)
		CP mortality	2,516	106,409	12.27	1.05 (0.99, 1.11)
		AMI incidence	2,189	105,695	12.28	1.09 (1.03, 1.15)
		Stroke incidence	1,507	105,981	12.30	1.01 (0.93, 1.08)
PM2.5	$\mu g/m^3$	All cause mortality	4,783	98,426	9.29	1.59 (1.53, 1.64)
		CP mortality	2,296	98,426	9.30	1.57 (1.50, 1.65)
		AMI incidence	1,966	97,750	9.31	1.70 (1.62, 1.79)
		Stroke incidence	1,379	98,017	9.31	1.73 (1.64, 1.83)
PM10	$\mu g/m^3$	All cause mortality	3,525	68,957	16.56	0.98 (0.92, 1.04)
		CP mortality	1,739	68,957	16.58	1.00 (0.92, 1.08)
		AMI incidence	1,460	68,477	16.60	1.04 (0.95, 1.12)
		Stroke incidence	1,040	68,671	16.61	1.03 (0.93, 1.13)
NO_2	ppb	All cause mortality	922	16,636	17.11	1.12 (1.01, 1.23)
		CP mortality	438	16,636	17.14	1.05 (0.89, 1.21)
		AMI incidence	375	16,515	17.05	1.25 (1.09, 1.42)
		Stroke incidence	280	16,557	17.08	1.28 (1.09, 1.47)
NO_X	ppb	All cause mortality	909	16,497	48.70	1.03 (0.93, 1.12)
		CP mortality	433	16,497	48.71	0.92 (0.78, 1.05)
		AMI incidence	371	16,377	48.66	1.12 (0.97, 1.26)
		Stroke incidence	279	16,420	48.69	1.14 (0.97, 1.31)
SO_2	ppb	All cause mortality	1,091	20,785	0.86	1.03 (0.95, 1.11)
		CP mortality	542	20,785	0.86	1.05 (0.93, 1.16)
		AMI incidence	453	20,648	0.86	1.02 (0.90, 1.14)
		Stroke incidence	312	20,693	0.86	1.10 (0.95, 1.25)
CO	ppm	All cause mortality	601	10,623	0.44	1.21 (1.11, 1.30)
		CP mortality	287	10,623	0.44	1.06 (0.91, 1.21)
		AMI incidence	227	10,542	0.44	1.36 (1.21, 1.51)
		Stroke incidence	177	10,567	0.44	1.29 (1.12, 1.46)

Table 11b. Hazard ratios for average air pollutant exposures (1995-2002) in relationship to disease outcomes for nonmovers in CTS cohort based on pollutant interquartile ranges

Pollutant	Units	Outcome	# events	N	IQR	HR (95% CI)
Ozone	ppb	All cause mortality	3,229	77,942	12.03	1.00 (0.95, 1.05)
		CP mortality	1,504	77,942	12.05	1.06 (0.99, 1.14)
		AMI incidence	1,479	77,434	12.08	1.09 (1.02, 1.17)
		Stroke incidence	1,020	77,647	12.09	1.05 (0.96, 1.14)
PM2.5	$\mu g/m^3$	All cause mortality	2,980	72,152	9.32	1.63 (1.56, 1.69)
		CP mortality	1,386	72,152	9.34	1.60 (1.51, 1.70)
		AMI incidence	1,332	71,667	9.34	1.66 (1.56, 1.76)
		Stroke incidence	935	71,871	9.35	1.63 (1.51, 1.76)
PM10	$\mu g/m^3$	All cause mortality	2,131	50,256	16.52	1.00 (0.93, 1.07)
		CP mortality	1,025	50,256	16.55	1.03 (0.93, 1.13)
		AMI incidence	953	49,907	16.57	1.01 (0.91, 1.12)
		Stroke incidence	684	50,055	16.58	1.06 (0.94, 1.19)
NO_2	ppb	All cause mortality	558	12,023	17.04	1.11 (0.97, 1.25)
		CP mortality	260	12,023	17.07	1.03 (0.83, 1.24)
		AMI incidence	247	11,934	17.03	1.14 (0.93, 1.34)
		Stroke incidence	170	11,968	17.03	1.28 (1.03, 1.53)
NO_X	ppb	All cause mortality	549	11,923	48.78	1.02 (0.90, 1.15)
		CP mortality	258	11,923	48.83	0.90 (0.72, 1.09)
		AMI incidence	245	11,835	48.77	1.00 (0.82, 1.18)
		Stroke incidence	169	11,869	48.77	1.14 (0.93, 1.36)
SO_2	ppb	All cause mortality	667	15,210	0.85	1.10 (1.00, 1.20)
		CP mortality	319	15,210	0.85	1.18 (1.03, 1.33)
		AMI incidence	294	15,100	0.85	1.03 (0.88, 1.18)
		Stroke incidence	210	15,143	0.85	1.13 (0.95, 1.31)
CO	ppm	All cause mortality	356	7,604	0.44	1.21 (1.08, 1.34)
		CP mortality	168	7,604	0.44	1.16 (0.96, 1.36)
		AMI incidence	154	7,545	0.44	1.24 (1.05, 1.43)
		Stroke incidence	104	7,567	0.44	1.23 (0.99, 1.47)

Table 11c. Hazard ratios for average air pollutant exposures (1995-2002) in relationship to disease outcomes for movers in the CTS cohort based on pollutant interquartile ranges

Pollutant	Units	Outcome	# events	N	Inter-quartile range	HR (95% CI)
Ozone	ppb	All cause mortality	1,990	28,467	11.77	0.99 (0.93, 1.06)
		CP mortality	1,012	28,467	11.79	1.05 (0.96, 1.14)
		AMI incidence	710	28,261	11.81	1.09 (0.98, 1.19)
		Stroke incidence	487	28,334	11.85	0.92 (0.79, 1.06)
PM2.5	$\mu g/m^3$	All cause mortality	1,803	26,274	8.85	1.57 (1.49, 1.66)
		CP mortality	910	26,274	8.87	1.59 (1.47, 1.71)
		AMI incidence	634	26,083	8.88	1.83 (1.69, 1.97)
		Stroke incidence	444	26,146	8.88	2.00 (1.83, 2.16)
PM10	$\mu g/m^3$	All cause mortality	1,394	18,701	16.91	1.00 (0.90, 1.09)
		CP mortality	714	18,701	16.92	1.03 (0.90, 1.16)
		AMI incidence	507	18,570	16.95	1.12 (0.96, 1.27)
		Stroke incidence	356	18,616	16.98	1.02 (0.84, 1.21)
NO_2	ppb	All cause mortality	364	4,613	17.62	1.12 (0.93, 1.30)
		CP mortality	178	4,613	17.58	1.04 (0.76, 1.32)
		AMI incidence	128	4,581	17.52	1.50 (1.20, 1.79)
		Stroke incidence	110	4,589	17.58	1.25 (0.92, 1.58)
NO_X	ppb	All cause mortality	360	4,574	49.86	1.04 (0.88, 1.20)
		CP mortality	175	4,574	49.85	0.95 (0.70, 1.20)
		AMI incidence	126	4,542	49.75	1.34 (1.08, 1.60)
		Stroke incidence	110	4,551	49.76	1.14 (0.86, 1.43)
SO_2	ppb	All cause mortality	424	5,575	0.90	0.97 (0.85, 1.10)
		CP mortality	223	5,575	0.90	0.98 (0.79, 1.16)
		AMI incidence	159	5,548	0.91	0.98 (0.78, 1.19)
		Stroke incidence	102	5,550	0.90	1.00 (0.74, 1.25)
CO	ppm	All cause mortality	245	3,019	0.53	1.10 (0.92, 1.28)
		CP mortality	119	3,019	0.53	0.89 (0.59, 1.18)
		AMI incidence	76	2,997	0.52	1.46 (1.15, 1.77)
		Stroke incidence	73	3,000	0.52	1.41 (1.13, 1.70)

Table 12a. Hazard ratios for average air pollutant exposures (1988-2002) in relationship to disease outcomes for full CTS cohort based on pollutant interquartile ranges

Pollutant	Units	Outcome	# events	N	Inter-quartile range	HR (95% CI)
Ozone	ppb	All cause mortality	5,219	106,409	16.41	1.16 (1.12, 1.20)
		CP mortality	2,516	106,409	16.58	1.20 (1.14, 1.26)
		AMI incidence	2,189	105,695	16.64	1.29 (1.23, 1.36)
		Stroke incidence	1,507	105,981	16.63	1.19 (1.11, 1.27)
PM2.5	$\mu g/m^3$	All cause mortality	4,783	98,426	8.45	1.54 (1.49, 1.58)
		CP mortality	2,296	98,426	8.49	1.52 (1.46, 1.59)
		AMI incidence	1,966	97,750	8.48	1.63 (1.56, 1.70)
		Stroke incidence	1,379	98,017	8.48	1.57 (1.48, 1.66)
PM10	$\mu g/m^3$	All cause mortality	3,525	68,957	14.78	1.12 (1.07, 1.16)
		CP mortality	1,739	68,957	14.82	1.14 (1.07, 1.20)
		AMI incidence	1,460	68,477	14.85	1.20 (1.13, 1.27)
		Stroke incidence	1,040	68,671	14.89	1.15 (1.06, 1.23)
NO_2	ppb	All cause mortality	922	16,636	18.39	1.12 (1.03, 1.22)
		CP mortality	438	16,636	18.42	1.08 (0.94, 1.23)
		AMI incidence	375	16,515	18.39	1.26 (1.11, 1.42)
		Stroke incidence	280	16,557	18.35	1.23 (1.05, 1.41)
NO_X	ppb	All cause mortality	909	16,497	49.10	1.06 (0.97, 1.15)
		CP mortality	433	16,497	49.12	0.99 (0.86, 1.12)
		AMI incidence	371	16,377	48.97	1.15 (1.02, 1.28)
		Stroke incidence	279	16,420	48.96	1.11 (0.95, 1.26)
SO_2	ppb	All cause mortality	1,091	20,785	0.94	1.00 (0.92, 1.07)
		CP mortality	542	20,785	0.94	0.98 (0.87, 1.09)
		AMI incidence	453	20,648	0.95	0.99 (0.87, 1.11)
		Stroke incidence	312	20,693	0.95	1.08 (0.93, 1.22)
CO	ppm	All cause mortality	601	10,623	0.54	1.17 (1.07, 1.26)
		CP mortality	287	10,623	0.53	1.05 (0.90, 1.19)
		AMI incidence	227	10,542	0.53	1.30 (1.15, 1.44)
		Stroke incidence	177	10,567	0.53	1.26 (1.09, 1.43)

Table 12b. Hazard ratios for average air pollutant exposures (1988-2002) in relationship to disease outcomes for nonmovers in CTS cohort based on pollutant interquartile ranges

Pollutant	Units	Outcome	# events	N	IQR	HR (95% CI)
Ozone	ppb	All cause mortality	3,229	77,942	17.13	1.18 (1.13, 1.24)
		CP mortality	1,504	77,942	17.21	1.24 (1.16, 1.32)
		AMI incidence	1,479	77,434	17.26	1.32 (1.24, 1.41)
		Stroke incidence	1,020	44,647	17.25	1.23 (1.13, 1.33)
PM2.5	$\mu g/m^3$	All cause mortality	2,908	72,152	8.44	1.57 (1.51, 1.63)
		CP mortality	1,386	72,152	8.45	1.56 (1.47, 1.65)
		AMI incidence	1,332	71,667	8.47	1.63 (1.54, 1.72)
		Stroke incidence	935	71,871	8.47	1.54 (1.43, 1.64)
PM10	$\mu g/m^3$	All cause mortality	2,131	50,256	14.74	1.13 (1.08, 1.19)
		CP mortality	1,025	50,256	14.73	1.16 (1.07, 1.24)
		AMI incidence	953	49,907	14.74	1.19 (1.10, 1.28)
		Stroke incidence	684	50,055	14.73	1.18 (1.08, 1.29)
NO_2	ppb	All cause mortality	558	12,023	18.40	1.12 (0.99, 1.25)
		CP mortality	260	12,023	18.44	1.08 (0.89, 1.27)
		AMI incidence	247	11,934	18.40	1.17 (0.98, 1.36)
		Stroke incidence	170	11,968	18.39	1.24 (1.01, 1.47)
NO_X	ppb	All cause mortality	549	11,923	49.60	1.05 (0.94, 1.17)
		CP mortality	258	11,923	49.62	0.98 (0.81, 1.15)
		AMI incidence	245	11,835	49.48	1.06 (0.89, 1.22)
		Stroke incidence	169	11,869	49.37	1.12 (0.91, 1.32)
SO_2	ppb	All cause mortality	667	15,210	0.92	1.04 (0.94, 1.13)
		CP mortality	319	15,210	0.92	1.07 (0.93, 1.21)
		AMI incidence	294	15,100	0.93	1.02 (0.87, 1.16)
		Stroke incidence	210	15,143	0.92	1.09 (0.91, 1.26)
CO	ppm	All cause mortality	356	7,604	0.53	1.17 (1.04, 1.29)
		CP mortality	168	7,604	0.53	1.19 (0.99, 1.38)
		AMI incidence	154	7,545	0.53	1.28 (1.09, 1.46)
		Stroke incidence	104	7,567	0.53	1.21 (0.98, 1.44)

Table 12c. Hazard ratios for average air pollutant exposures (1988-2002) in relationship to disease outcomes for movers in the CTS cohort based on pollutant interquartile ranges

Pollutant	Units	Outcome	# events	N	IQR	HR (95% CI)
Ozone	ppb	All cause mortality	1,990	28,467	14.72	1.11 (1.04, 1.17)
		CP mortality	1,012	28,467	14.91	1.14 (1.05, 1.23)
		AMI incidence	710	28,261	14.99	1.24 (1.14, 1.34)
		Stroke incidence	487	28,334	15.01	1.11 (0.98, 1.23)
PM2.5	$\mu g/m^3$	All cause mortality	1,803	26,274	8.10	1.47 (1.40, 1.55)
		CP mortality	910	26,274	8.11	1.46 (1.35, 1.57)
		AMI incidence	634	26,083	8.11	1.60 (1.48, 1.73)
		Stroke incidence	455	26,146	8.10	1.65 (1.50, 1.81)
PM10	$\mu g/m^3$	All cause mortality	1,394	18,701	16.72	1.10 (1.02, 1.19)
		CP mortality	714	18,701	16.70	1.13 (1.01, 1.26)
		AMI incidence	507	18,570	16.69	1.23 (1.09, 1.37)
		Stroke incidence	356	18,616	16.73	1.10 (0.93, 1.27)
NO_2	ppb	All cause mortality	364	4,613	18.98	1.05 (0.88, 1.21)
		CP mortality	178	4,613	18.96	1.01 (0.76, 1.26)
		AMI incidence	128	4,581	18.87	1.38 (1.11, 1.66)
		Stroke incidence	110	4,589	18.95	1.20 (0.90, 1.49)
NO_X	ppb	All cause mortality	360	4,574	47.59	1.02 (0.88, 1.16)
		CP mortality	175	4,574	47.57	0.97 (0.76, 1.18)
		AMI incidence	126	4,542	47.41	1.29 (1.06, 1.52)
		Stroke incidence	110	4,551	47.44	1.09 (0.84, 1.34)
SO_2	ppb	All cause mortality	424	5,575	1.15	0.94 (0.79, 1.08)
		CP mortality	223	5,575	1.15	0.89 (0.68, 1.10)
		AMI incidence	159	5,548	1.15	0.98 (0.74, 1.22)
		Stroke incidence	102	5,550	1.15	1.04 (0.73, 1.34)
CO	ppm	All cause mortality	245	3,019	0.55	1.00 (0.84, 1.17)
		CP mortality	119	3,019	0.55	0.79 (0.53, 1.06)
		AMI incidence	76	2,997	0.55	1.19 (0.89, 1.49)
		Stroke incidence	73	3,000	0.55	1.39 (1.10, 1.69)

Table 13. Hazard ratios per interquartile range for traffic, road and vehicle metrics using full regression models, adjusted for age, race, and covariates*

Variable**	Outcome	# events	N	IQR	HR (95% CI)
Distance to highway	All cause mortality	5,370	109,039	1,821	0.99 (0.98, 1.01)
	CP mortality	2,601	109,039	1,816	1.01 (0.98, 1.03)
	AMI incidence	2,251	108,298	1,816	0.97 (0.94, 1.00)
	Stroke incidence	1,554	108,598	1,813	1.00 (0.97, 1.03)
Traffic density 150m	All cause mortality	5,370	109,039	1,430	1.00 (0.99, 1.01)
	CP mortality	2,601	109,039	1,444	1.00 (0.98, 1.02)
	AMI incidence	2,251	108,298	1,451	0.99 (0.97, 1.01)
	Stroke incidence	1,554	108,598	1,453	1.00 (0.97, 1.02)
Traffic density 300m	All cause mortality	5,370	109,039	1,866	1.01 (0.99, 1.03)
	CP mortality	2,601	109,039	1,873	1.02 (0.99, 1.04)
	AMI incidence	2,251	108,298	1,877	0.99 (0.96, 1.02)
	Stroke incidence	1,554	108,598	1,876	1.00 (0.97, 1.03)
Road Density	All cause mortality	5,370	109,018	422	1.03 (1.00, 1.07)
	CP mortality	2,601	109,018	422	1.00 (0.95, 1.05)
	AMI incidence	2,251	108,277	423	1.04 (0.99, 1.09)
	Stroke incidence	1,554	108,577	423	1.07 (1.01, 1.13)
Vehicle Density	All cause mortality	5,370	109,039	3,715	1.02 (0.99, 1.05)
	CP mortality	2,601	109,039	3,712	1.02 (0.97, 1.06)
	AMI incidence	2,251	109,298	3,709	1.00 (0.94, 1.05)
	Stroke incidence	1,554	108,598	3,707	1.05 (0.99, 1.11)

* Adjusted for smoking status, total pack years, BMI, marital status, alcohol consumption, SHS home exposure, dietary fat, dietary fiber, dietary calories, physical activity, menopausal status, HT use; and contextual variables (income, income inequality, education, population size, racial composition, unemployment). Regressions for outcomes other than all-cause mortality also included family history of AMI, family history of stroke, blood pressure medication and aspirin use.

** Traffic variable definitions:

Distance to highway= Proximity of residence to a "major" highway, in meters. (Limited to within 20km.) Missing data (n=180) changed to 49999.

Traffic density 150m = Vehicle Miles Traveled within 150 meters of a residence using conflated TeleAtlas 2005q2 centerlines linked to HPMS 2000. Missing values (n=46,909) set to minimum non-zero value (0.10442).

Traffic density 300m = Vehicle Miles Traveled within 300 meters of a residence using conflated TeleAtlas 2005q2 centerlines linked to HPMS 2000. (Normalized to 150m values.) Missing values (n=19,700) set to minimum non-zero value (0.00339).

Road Density = Meters of roads (based on TeleAtlas/Dynamap road data) within 150 meters of a residence

Vehicle Density = 2000 Census Block group count of aggregate number of vehicles available from occupied housing units.

Note: Traffic metrics were also calculated with missing values excluded and the regression results were similar (results not shown). Also, analyses using traffic density metrics calculated from CalTrans traffic data were similar to those in the above table (results not shown).

The remaining tables present the results of several sensitivity analyses. Specifically, we examined the effects of using only measured, and not estimated, values of PM2.5. Measured values were only available for 1999 through 2002, which covers half of the follow-up period; the numbers of observed events in all four outcome categories were somewhat more than half of those for the corresponding outcomes for the full follow-up period, consistent with the aging of the cohort. However, this reduction in power may be offset to some extent by a reduction in measurement error, since only measured pollutant values were used. Second, we examined multi-pollutant models for all of the major pollutants. Third, we examined two-pollutant models for PM2.5 and ozone, with both limited only to the period 1999 through 2002. Fourth, we examined the impact of using only third quarter ozone values, which correspond to the highest levels of personal ozone exposure during any given year not only because ozone levels are elevated during these months, but also because people are more likely to be outdoors and have their windows open. Finally, we present the results of our examination of the data for spatial autocorrelation.

Table 14 summarizes the results for PM2.5 using the 1999-2002 exposure and follow-up period, during which only measured values of PM2.5 were used. Associations were observed between PM2.5 and all four outcomes. Specifically, for the full cohort, a one $\mu g/m^3$ increase in PM2.5 exposure was associated with a 0.8% change in total mortality, a 1.3% change in cardiopulmonary mortality, and 1.4% increases in the incidence of both stroke and AMI. Similar effect estimates were observed for the subcohort of nonmovers. Much greater effects were observed in the sub-cohort of movers, for whom a one $\mu g/m^3$ was associated with 2.3% changes in total mortality and AMI incidence, a 2.9% change in cardiopulmonary mortality, and a 3% change in stroke incidence.

Table 15 summarizes the results from two-pollutant models, using the exposure and follow-up data for 1995 through 2002. When both PM2.5 and either ozone or NO_2 are included in the model, the effects of PM2.5 on cardiopulmonary mortality and AMI incidence are slightly greater than those of the single-pollutant models for PM2.5, while the associations of the gases with these outcomes are attenuated to the point where they even appear to be slightly protective (i.e., the point estimates and confidence intervals are all less than unity). Adding CO to the model increases the magnitude of the association between PM2.5 and cardiopulmonary mortality and decreases the PM2.5 coefficient to nonsignificance in relation to AMI. CO was strongly associated with AMI incidence, regardless of which other pollutant was included in the model. It should be noted that, because our buffer zones and extent of interpolation for CO and NO_2 were narrowly circumscribed (see Table 1), the numbers of participants and events included in two-pollutant models involving these gases were far smaller than those involving only PM2.5 and ozone. Therefore, the resulting estimates (especially for CO) tended to be much less precise than for the pollutants with broader catchment areas.

Table 14. Hazard ratios for PM2.5, per μg/m^3, using only measured (not reconstructed) data from 1999-2002 for the full cohort, nonmovers, and movers

Total Cohort

	Units	Mean (SD)	Range	Outcome	# events	N	HR (95% CI)
PM2.5	μg/m^3	17.24 (4.94)	4.86-29.76	All cause mortality	2,853	93,286	1.008 (1.000, 1.016)
				CP mortality	1,410	93,286	1.013 (1.002, 1.025)
				AMI incidence	1,160	92,184	1.014 (1.002, 1.027)
				Stroke incidence	811	92,604	1.014 (1.000, 1.029)

Nonmovers

Pollutant	Units	Mean (SD)	Range	Outcome	# events	N	HR (95% CI)
PM2.5	μg/m^3	17.20 (4.94)	5.05-29.76	All cause mortality	1,824	69,439	1.007 (0.997, 1.017)
				CP mortality	880	69,439	1.013 (0.998, 1.027)
				AMI incidence	806	68,633	1.013 (0.998, 1.028)
				Stroke incidence	567	68,960	1.014 (0.996, 1.032)

Movers

Pollutant	Units	Mean (SD)	Range	Outcome	# events	N	HR (95% CI)
PM2.5	μg/m^3	17.03 (4.82)	4.46-40.17	All cause mortality	1,029	23,847	1.023 (1.009, 1.037)
				CP mortality	530	23,847	1.029 (1.009, 1.049)
				AMI incidence	353	23,550	1.023 (0.999, 1.047)
				Stroke incidence	244	23,644	1.030 (1.001, 1.059)

Table 15. Hazard ratios for cardiopulmonary (CP) mortality and AMI incidence per unit pollutant in two-pollutant models using exposures from 1995-2002 for the full cohort

Outcome	# events	N	Pollutants in model	HR (95% CI)
CP mortality	2,265	97,283	PM2.5	1.063 (1.053, 1.073)
			Ozone	0.985 (0.979, 0.991)
AMI incidence	1,948	96,621	PM2.5	1.068 (1.058, 1.079)
			Ozone	0.988 (0.981, 0.994)
CP mortality	422	16,064	PM2.5	1.071 (1.032, 1.111)
			NO_2	0.976 (0.959, 0.994)
AMI incidence	356	15,942	PM2.5	1.094 (1.050, 1.140)
			NO_2	0.979 (0.960, 0.999)
CP mortality	268	10,048	PM2.5	1.084 (1.042, 1.127)
			CO	0.577 (0.340, 0.978)
AMI incidence	202	9,965	PM2.5	1.026 (0.983, 1.071)
			CO	2.068 (1.165, 3.669)
CP mortality	287	10,623	CO	1.135 (0.810, 1.590)
			Ozone	1.007 (0.992, 1.021)
AMI incidence	227	10,542	CO	2.035 (1.452, 2.852)
			Ozone	0.989 (0.973, 1.006)
CP mortality	241	8,891	CO	0.983 (0.549, 1.761)
			NO_2	1.005 (0.984, 1.027)
AMI incidence	195	8,822	CO	3.754 (1.822, 7.732)
			NO_2	0.977 (0.952, 1.003)
CP mortality	438	16,636	Ozone	0.995 (0.983, 1.007)
			NO_2	1.005 (0.994, 1.015)
AMI incidence	375	16,515	Ozone	0.996 (0.983, 1.009)
			NO_2	1.015 (1.004, 1.026)

Table 16 presents the results of two-pollutant models for cardiopulmonary mortality and AMI incidence in the full cohort for the period 1999 through 2002. This analysis was intended to examine the robustness of the estimates for PM2.5 during the period in which only measured values for this pollutant were used to develop exposure estimates. Including ozone in the model increased the PM2.5 coefficients from 1.013 to 1.027 for cardiopulmonary mortality, while the coefficient for AMI incidence remained roughly the same (1.014 in the single-pollutant model and 1.016 in the model with ozone). The ozone coefficient and confidence interval for cardiopulmonary mortality were less than unity, while there was no relation between ozone and AMI incidence in the two-pollutant model. (In single-pollutant models for ozone measured from 1999 through 2002, the ozone hazard ratios per ppb were 0.994 (95% CI = 0.988, 1.001) for cardiopulmonary mortality and 1.000 (95% CI = 0.993, 1.007) for AMI incidence.)

Table 17 summarizes the results for models in which long-term exposure to ozone was limited to the third quarter values for this pollutant for the period 1995 through 2002. In single-pollutant models, ozone was associated with all four outcomes. In models using the full-year average (Table 10a), ozone was associated only with cardiopulmonary mortality (in the simple model) and with AMI incidence. For two-pollutant models that include (annual) PM2.5, however, the ozone coefficients all diminish in magnitude and become nonsignificant.

Table 16. Hazard ratios in multi-pollutant models measuring associations between average air pollutant exposures (1999-2002) and cardiopulmonary (CP) mortality and AMI incidence among the full cohort

				Adjusted for age and race	**Adjusted for age, race, and covariates***
Outcome	**# events**	**N**	**Pollutants in model**	**HR (95% CI)**	**HR (95% CI)**
CP mortality	1,385	92,116	PM2.5	1.022 (1.010, 1.035)	1.027 (1.013, 1.040)
			Ozone	0.992 (0.985, 0.999)	0.987 (0.979, 0.994)
AMI incidence	1,151	91,029	PM2.5	1.015 (1.001, 1.028)	1.016 (1.002, 1.031)
			Ozone	1.001 (0.994, 1.009)	0.995 (0.987, 1.004)

Covariates in models include: smoking status, total pack years (current and former smokers), BMI, marital status, alcohol consumption, SHS exposure in the home, dietary fat, dietary fiber, dietary calories, physical activity, menopausal status, hormone therapy, family history of AMI or stroke, use of blood pressure medication, aspirin use, and contextual variables (1990 Census variables by block group) consisting of median household income, income inequality (percent in poverty), education (percent with bachelor degree and above), population size (total population), racial/ethnic composition (percent Black, White, Hispanic),and unemployment (percent unemployed age 16+).

Table 17. Results using third-quarter ozone in single-and multi-pollutant models using exposures from 1995-2002 for the full cohort

Outcome	# events	N	Pollutants in model	HR (95% CI)
All cause mortality	5,015	106,002	Ozone (3rd quarter)	1.004 (1.002, 1.006)
CP mortality	2,417	106,002	Ozone (3rd quarter)	1.006 (1.003, 1.008)
AMI incidence	2,096	105,229	Ozone (3rd quarter)	1.009 (1.006, 1.012)
Stroke incidence	1,434	105,529	Ozone (3rd quarter)	1.004 (1.001, 1.007)
All cause mortality	4,553	96,926	PM2.5	1.050 (1.042, 1.057)
			Ozone (3rd quarter)	0.995 (0.993, 0.997)
CP mortality	2,178	96,926	PM2.5	1.043 (1.032, 1.054)
			Ozone (3rd quarter)	0.998 (0.994, 1.001)
AMI incidence	1,893	96,211	PM2.5	1.044 (1.032, 1.056)
			Ozone (3rd quarter)	1.001 (0.998, 1.005)
Stroke incidence	1,305	96,488	PM2.5	1.059 (1.045, 1.074)
			Ozone (3rd quarter)	0.993 (0.988, 0.997)

Finally, we had somewhat mixed results from the spatial autocorrelation modeling. In general we found little evidence of spatial autocorrelation in our data, once the other nonpollutant covariates were entered into the models. The random effects model worked well for some pollutants (e.g., ozone) but not so well for others (notably PM2.5). Table 18 compares results from the traditional proportional hazards model (Table 11a), which make no allowance for spatial dependencies, to two proportional hazards models with additional random effects. These models have the potential to control for residual confounding that may be spatially related, that is, confounding by unknown risk factors that cluster spatially, above and beyond what can be accounted for by the covariates in the model. The two additional models in Table 18 represent different degrees of spatial organization on the county and zip code level, as described in the Methods section: (i) two-level clusters (county and zip code) allowing women within these areas to share a common risk, independent of those in nearby areas, and (ii) two-level clusters that are dependent on the risks in adjacent clusters.

The random effects models were fit with software that had been previously applied to the ACS-CPS II analysis, but occasionally returned unusual results with the CTS data. In particular, models run with exposure data averaged from 1988 often failed to converge (results not shown). Results from the random effects models using the 1995-2002 data (Table 18) were generally consistent with the previous results, indicating little evidence for residual spatial autocorrelation. Exceptions were some of the results for PM 2.5 and CO, for which the spatial clustering models produced anomalous results (e.g., impossibly small standard errors). The difficulties with the modeled PM2.5 may be due to the sparseness of observations in some zip codes, though there may be other factors involved. For example, if some participants were assigned the same values for Census-

level or contextual variables, then the validity of the underlying assumption that all observations within the zip codes are independent may be questionable. Additional analyses will be required to identify the circumstances that produce such results.

Table 18. Hazard ratios for average air pollutant exposures (1995-2002) in relationship to disease outcomes for full CTS cohort based on pollutant interquartile ranges, with random effects models for spatial dependence

		No spatial random effects	Two-level (zip and county) conditionally independent	Two-level (zip and county) conditionally dependent
Pollutant	**Outcome**	**HR (95% CI)**	**HR (95% CI)**	**HR (95% CI)**
Ozone	All cause mortality	1.00 (0.96, 1.04)	1.00 (0.96, 1.04)	1.00 (0.96, 1.05)
	CP mortality	1.05 (0.99, 1.11)	1.05 (0.99, 1.11)	1.06 (1.00, 1.12)
	AMI incidence	1.09 (1.03, 1.15)	1.09 (1.03, 1.16)	1.09 (1.03, 1.16)
	Stroke incidence	1.01 (0.93, 1.08)	1.01 (0.94, 1.09)	1.01 (0.94, 1.09)
PM2.5	All cause mortality	1.59 (1.53, 1.64)	1.74 (1.50, 2.00)	3.11 (2.58, 3.75)
	CP mortality	1.57 (1.50, 1.65)	1.60 (1.31, 1.97)	1.60 (1.31, 1.97)
	AMI incidence	1.70 (1.62, 1.79)	*unstable results*	*unstable results*
	Stroke incidence	1.73 (1.64, 1.83)	8.50 (7.30, 9.90)	15.4 (13.3, 17.7)
PM10	All cause mortality	0.98 (0.92, 1.04)	0.98 (0.93, 1.04)	0.96 (0.90, 1.03)
	CP mortality	1.00 (0.92, 1.08)	1.00 (0.92, 1.08)	1.00 (0.92, 1.08)
	AMI incidence	1.04 (0.95, 1.12)	1.03 (0.95, 1.13)	1.03 (0.95, 1.13)
	Stroke incidence	1.03 (0.93, 1.13)	1.03 (0.93, 1.14)	1.03 (0.93, 1.14)
NO_2	All cause mortality	1.12 (1.01, 1.23)	1.22 (1.08, 1.38)	1.49 (1.27, 1.74)
	CP mortality	1.05 (0.89, 1.21)	1.08 (0.93, 1.27)	1.08 (0.93, 1.27)
	AMI incidence	1.25 (1.09, 1.42)	1.27 (1.08, 1.50)	1.27 (1.08, 1.50)
	Stroke incidence	1.28 (1.09, 1.47)	1.28 (1.06, 1.55)	1.28 (1.06, 1.55)
NO_X	All cause mortality	1.03 (0.93, 1.12)	1.03 (0.94, 1.13)	1.05 (0.95, 1.17)
	CP mortality	0.92 (0.78, 1.05)	0.94 (0.82, 1.07)	0.94 (0.82, 1.07)
	AMI incidence	1.12 (0.97, 1.26)	1.13 (0.98, 1.31)	1.13 (0.98, 1.31)
	Stroke incidence	1.14 (0.97, 1.31)	1.13 (0.96, 1.34)	1.13 (0.96, 1.34)
SO_2	All cause mortality	1.03 (0.95, 1.11)	1.02 (0.95, 1.11)	1.02 (0.95, 1.11)
	CP mortality	1.05 (0.93, 1.16)	1.04 (0.93, 1.16)	1.04 (0.93, 1.16)
	AMI incidence	1.02 (0.90, 1.14)	1.00 (0.89, 1.13)	1.00 (0.89, 1.13)
	Stroke incidence	1.10 (0.95, 1.25)	1.10 (0.95, 1.27)	1.10 (0.95, 1.27)
CO	All cause mortality	1.21 (1.11, 1.30)	0.95 (0.72, 1.25)	0.95 (0.72, 1.25)
	CP mortality	1.06 (0.91, 1.21)	0.73 (0.47, 1.15)	0.73 (0.47, 1.15)
	AMI incidence	1.36 (1.21, 1.51)	1.34 (1.16, 1.55)	1.34 (1.16, 1.55)
	Stroke incidence	1.29 (1.12, 1.46)	1.31 (1.11, 1.56)	1.31 (1.11, 1.56)

Discussion

In an ongoing cohort study of over 100,000 female participants in the California Teachers Study, we were able to develop estimates of long-term air pollution exposure at the subjects' residences and to examine associations between these exposure estimates and several serious outcomes associated with circulatory diseases. This represents a significant improvement over other air pollution cohort studies that used fixed-site monitors for a limited number of years as the basis for pollution metrics. We identified a number of statistically significant associations between long-term exposures and total mortality, cardiopulmonary mortality, incidence of acute myocardial infarction and stroke. In addition, we examined the potential impacts of several traffic metrics on these outcomes.

The low prevalence of active smoking in this cohort (5% at baseline), in combination with previously collected data on household exposure to SHS, allowed for a closer examination of the impact of air pollution exposures during the follow-up period than in other investigations in which active smoking had to be controlled for statistically (e.g., Pope et al. 1995, 22% active smokers; Dockery et al. 33 – 40% smokers, depending on the city). Nationally, the age-adjusted prevalence of active smoking among women in the U.S. was 21% in 2000 (Centers for Disease Control 2002). In California, the age-adjusted smoking prevalence among women was 13.6 ($\pm$ 2.3%) in 1995 and 14.4% ($\pm$ 1.6%) in 2000, indicating that even in California, the CTS participants were substantially less likely to be smokers than women in the general population (Centers for Disease Control 1996; 2001). Despite the low prevalence of active smoking, we found that among smokers, there was approximately a 24% increase in risk of dying from cardiovascular disease during the follow-up period (Table 7).

Household exposure to SHS was present at baseline in approximately half the population in this analysis, which represented a substantially higher proportion than the 1996 statewide prevalence of smoking among men (19.8%) (California Department of Health Services, http://www.cstats.info). By 2002 the statewide male smoking prevalence had declined slightly (19.5%), with a more marked decline among men $\geq$ 65 years of age (10.9% to 8.3%), presumably due to increased mortality rates among older smokers. To the extent that the CTS participants' spouses' smoking behaviors were similar to those of the California male population, SHS exposures were likely to have been somewhat less common among cohort members in 2002 than at baseline. Nevertheless, reported household SHS exposure was associated with a 7.2% statistically significant increase in risk of cardiovascular mortality (Table 7). This estimate is at the low end of those reported in studies examining relationships between SHS and cardiovascular outcomes in the past decade (OEHHA/ARB 2005). As noted in the Results section, the hazard ratios for other risk factors in the regression models are virtually all in the expected directions, which provides a check on the internal validity of the data and modeling used in this analysis.

Except for the ACS-CPS II, this is the largest air pollution cohort study undertaken to date. A majority of the women were post-menopausal at baseline (in 1995), putting them at increased risk for the development of cardiovascular disease. A recent study linking air pollution to atherosclerosis reported that postmenopausal women appear most susceptible (Kunzli et al. 2005). Moreover, as noted in the Introduction,

unlike most other cohort studies, there was a relative uniformity of occupational status that did not necessitate controlling for the kinds of potentially toxic exposures that would be common in industrial environments. By comparison, the prevalence of inhalation occupational hazards ranged from 28 – 53% in the Harvard Six Cities study (Dockery et al. 1993). Thus, several CTS cohort characteristics allowed for a number of advantages in examining circulatory outcomes in relation to air pollution.

Re-analysis of the Harvard Six Cities and the ACS-CPS II investigations indicated that mortality effects were limited to those who did not have more than a high school education (HEI 2000), suggesting an interaction of pollution with factors linked with educational attainment and lower socioeconomic status. In contrast, most (if not all) of the CTS cohort participants had more than a high school education, yet we were nevertheless able to detect associations of PM2.5 and other pollutants with mortality and incidence of AMI and stroke. Unfortunately, none of the CTS questionnaires inquired about the members' education, income or potential occupational exposures, so we could not analyze this issue in greater depth.

For those women in the CTS cohort who were still actively employed, the database did not include school address information, so it is possible that there was additional exposure misclassification in this group with respect to traffic exposures (e.g., at work or during commuting). We had proposed a sensitivity analysis (Lipsett 2004) to examine the extent to which the absence of the school addresses (for those teachers still working) might contribute to exposure misclassification, and therefore affect the risk estimates. This would have involved stratification of the cohort by retirement status at baseline, comparing the coefficients for those who were retired throughout the follow-up period with those for the rest of the cohort, after adjusting for age differences. However, upon closer examination we found that the "retirement" question in the baseline survey did not actually allow for a clear categorization of retirement status (i.e., it asked only about the dates of school-based employment and could therefore have misclassified as retired those individuals who pursued other non-school careers). Therefore, we did not undertake this analysis and this issue remains unresolved. The only other study that has attempted to refine exposures using both residential and workplace addresses (AHSMOG – Abbey et al. 1999; Chen et al. 2005) may have other challenges with exposure assessment (see below).

Two large cohort studies in the U.S. have identified fine particulate matter as the most important air pollutant with respect to cardiopulmonary mortality in the U.S. (Dockery et al. 1993; Laden et al. 2006; Pope et al. 1995, 2002, 2004). However, examination of the effects of PM2.5 have found a range of effects ranging from null results (Enstrom 2005) to effect estimates greater than those in the Harvard Six Cities Study and the ACS CPS II (Jerrett et al. 2005; Chen et al. 2005). Therefore, we were especially interested in investigating associations between PM2.5 exposures and the outcomes of interest in the CTS cohort. The results for the PM2.5 exposure periods that included substantial numbers of imputed values from the dataset provided by ARB are strikingly higher than any previously published studies, except for those of Chen et al. (2005). With reference to a 10 $\mu g/m^3$ increase in long-term exposure to PM2.5, the risk estimates based on the datasets including the imputed values are on the order of 1.5 to about 1.8, depending on the outcome. For purposes of comparison, estimates from most of these other studies are in the range of 1.06 – 1.30, depending on the outcome. While

part of these relatively high risk estimates might be explained by our more complete (both spatially and temporally) exposure data and lower nondifferential measurement error relative to previous studies, the use of the estimated PM2.5 data may have resulted in upwardly biased risk estimates (see below). In contrast, our sensitivity analysis using only measured PM2.5 data from 1999 through 2002 resulted in elevated hazard ratios for the full cohort and movers that were more consistent with estimates from prior studies. In addition, the hazard ratio for cardiopulmonary mortality using measured PM2.5 (but not when using the reconstructed PM2.5 data) was greater than that for all-cause mortality, which is also consistent with results of other studies.

We have not had the resources in this investigation to fully explore the reasons for these discrepancies. However, the values of PM2.5 predicted from other metrics in the historical reconstruction (Blanchard and Tanenbaum 2005) are likely to have underestimated the variance and extreme values PM2.5. This could potentially have led to systematic underestimation of the mean and variance of exposure among the CTS cohort and therefore systematic overestimation of the per-unit risks of mortality and circulatory disease incidence associated with exposure. In addition, the sensitivity analyses of Blanchard and Tanenbaum (2005) showed greater estimation errors in the winter months and with certain types of monitors, suggesting that the measurement errors may be non-random based on location. Therefore, we believe that the PM2.5 risks based on the reconstructed values should be interpreted with caution, while the hazard ratios using the measured PM2.5 data are likely to represent the best risk estimates available in this dataset. We recommend that ARB undertake additional work on the historical PM2.5 database, using both simulations and real data, to gain a better understanding of the magnitude of the exposure measurement error introduced through the use of regression-based imputed fine particle data.

The investigation by Chen et al. (2005), who reported highly elevated risks of CHD mortality for women in the AHSMOG study, may be subject to the same caveat. As with other AHSMOG investigations, the exposure assessment in that study involved careful interpolation of monthly pollutant concentrations to the centroids of zip codes in which the subjects' home and work addresses were located. However, all PM2.5 data and approximately half of the PM10 data were predicted based on regression models, with predicted PM2.5 being based on airport visibility measurements and predicted PM10 values on total suspended particle measurements. As such, the report by Chen et al. (2005) may have had similar issues concerning systematic underestimation of exposure and overestimation of risks.

In the CTS investigation, however, even the exclusive use of measured PM2.5 data indicated elevated risks, suggesting that exposure to this pollutant is linked not only with all-cause and cardiopulmonary mortality, but also with incident AMI and stroke (Table 14). Increases of 10 $\mu g/m^3$ PM2.5 in long-term average exposures were associated with significant increases in all-cause (8%) and cardiopulmonary (13%) mortality, as well as in the incidence of AMI (14%) and stroke (14%). These effect estimates were not affected by the inclusion of ozone (both full-year and third-quarter-only estimates) in the models, suggesting that they are not likely due to confounding by this widespread oxidant gas (Tables 16 and 17). Inclusion of full-year ozone in the model increased the estimated risk of cardiopulmonary mortality associated with a 10 $\mu g/m^3$ increase in long-term average PM2.5 from 13% to 27%, while the change in risk for AMI

increased only slightly (14% to 16%). Interestingly, the PM2.5 hazard ratios for movers were substantially higher than for nonmovers using this measured PM2.5 dataset, despite the clustering of several factors that would theoretically result in lower group risks for circulatory events among the movers, such as younger age, a greater fraction of pre-/perimenopausal versus post-menopausal women, greater reported physical activity, and a lower percentage of family histories for AMI and stroke (Table 2).

While a number of time-series analyses have linked daily changes in pollutant values with both hospitalizations for myocardial infarction and stroke (von Klot et al. 2005; Wellinius et al. 2005), this is the first prospective investigation of long-term exposure to PM2.5 to have reported associations with incident AMI and stroke. These analyses were, by definition, limited only to those who did not report a history of AMI or stroke on the baseline questionnaire. While some of those who did not report a prior history of such events might have had subclinical disease or might even have experienced a silent myocardial infarction or stroke, there is no reason to think *a priori* that such misclassification of disease would be distributed differentially among those with greater exposures to PM2.5 or other pollutants. Therefore, it is unlikely that the inclusion of subjects with subclinical disease would have resulted in upwardly biased hazard ratios. As the variables that we created for AMI and stroke represented both first hospitalizations for these events as well as mortality (with unique subject identifiers allowing us to avoid double-counting), we were able to capture a more complete picture of incident disease than other studies that have examined either hospitalizations or mortality.

Prior cohort studies of particulate matter have examined mortality from cardiovascular causes, but have not specifically examined incidence of new cases, and therefore could not distinguish between precipitation of an acute event in someone who already had the disease and the development of disease in the first place. Time-series studies of cardiovascular hospitalizations and mortality capture such acute events. There is growing evidence linking PM exposure with a variety of factors that might aggravate pre-existing atherosclerotic disease (resulting in myocardial infarction or dysrhythmia), such as arterial vasoconstriction, endothelial injury and dysfunction, and decreased heart rate variability (Brook et al. 2004). Recent toxicological evidence demonstrates that chronic exposure to low levels of PM2.5 (six-month study average 15 $\mu g/m^3$) is associated with progression of atherosclerotic disease, as well as increased vasomotor tone and vascular inflammation (Sun et al. 2005). Progression of atherosclerotic disease in humans can be observed subclinically as increases in carotid arterial intima medial thickness (CIMT, which represents thickening of the arterial wall). Increased CIMT has been reported cross-sectionally in association with estimated residential annual mean concentrations of PM2.5 in 798 subjects in Los Angeles (Kunzli et al. 2005). The Multi-Ethnic Study of Atherosclerosis Air Pollution Study (MESA AIR) will be examining the relationship of PM2.5 to CIMT progression prospectively over the next few years (http://depts.washington.edu/mesaair/). Thus, while we could not investigate subclinical outcomes in the CTS, our finding that long-term exposure to PM2.5 is associated with incident cases of both AMI and stroke is supported by recent mechanistic research.

The associations that we identified may represent effects of short-term exposures, longer-term exposures, or both. However, we were not able to investigate the relevant

windows of exposure associated with these outcomes. This issue will need to be resolved in future research.

In single-pollutant models, the length of the exposure period appears to play an important role in the magnitude and strength of associations between PM10 and all adverse health outcomes. We observed no significant associations between PM10 and any outcome when the exposure period was limited to 1995 through 2002. In contrast, with the longer exposure period (1988-2002) we observed significant associations between PM10 and all-cause and cardiopulmonary mortality, with essentially no difference in the magnitude of these hazard ratios. The estimated risk of death during the follow-up period per 10 $\mu g/m^3$ PM10 (HR = 1.09, 95% CI 1.04-1.13) is comparable to that in the ACS CPS-II study, but is lower than that in the recent AHSMOG publication (Pope et al. 2002; Chen et al. 2005.) Although, the AHSMOG participants lived only in California, the AHSMOG exposure period began in the mid-1970s, when PM levels were substantially higher than during our study. In addition, as noted previously, about half of the AHSMOG PM10 values were derived by regression methods from another PM metric, and therefore may have been subject to measurement errors underestimating PM10 variance, and overestimating the relative risk of exposure.

For the 1988-2002 exposure period, the hazard ratios linking PM10 and the incidence of both AMI and stroke were slightly higher than those for the mortality outcomes, with the exception of stroke incidence among movers only. The number of events in this last category was the lowest of any during this exposure period (n=354, compared with 507 – 3,525 for all of the others), suggesting that statistical power might have played a role in these results. The stark difference in the results for both the morbidity and mortality outcomes between the two exposure periods defies easy explanation. Considering that there are many studies linking daily PM10 with cardiovascular morbidity and mortality (e.g., Analitis et al. 2006; von Klot 2005), it is somewhat surprising that we did not detect such a signal for the 1995-2002 period. The mean estimated exposure concentrations to which the cohort members were exposed for the entire period were higher than those for the approximately seven-year follow-up, suggesting that exposure to these higher concentrations may have been important. On the other hand, we did not have residential histories on any of the participants prior to 1995 and assumed, for purposes of this analysis, 100% residential stability at the baseline address during the prior period (1988-1995). This assumption would have introduced substantial measurement error, which *a priori* was likely to have been nondifferential with respect to the outcomes. Therefore, this would have produced a bias towards the null hypothesis of no effect for period 1988-1995, resulting in an underestimate of the relationship between the PM10 and the various adverse outcomes. Finally, these results suggest that prolonged exposures to PM10 may be necessary to elicit the effects observed in this investigation. Additional analyses of critical exposure windows of exposure may help elucidate the temporal nature of these relationships.

The relationships of ozone hazard ratios for the two exposure periods followed a similar pattern as those for PM10, with stronger associations when the full exposure period was used. For the period 1995-2002, the hazard ratios in single-pollutant models for ozone were positive and significant, after controlling for relevant covariates, for incidence of AMI, and positive but of marginal statistical significance for cardiopulmonary mortality, both for the whole cohort and for nonmovers. The estimated

hazard ratios were of considerable magnitude; e.g., a 10 ppb increment in the average ozone exposure over the follow-up period was associated with an 8% increase in the risk of AMI among the nonmovers. However, the magnitude of the hazard ratios and their statistical significance were markedly greater when 1988-2002 was used as the exposure period. For instance, among nonmovers a 10 ppb increment in the average ozone exposure over the follow-up period was associated with 16% increase in the risk of AMI – twice the estimated risk compared with the 1995-2002 period. In addition, using the full exposure period, ozone was strongly associated with all four outcomes (except for stroke among movers). These hazard ratios were virtually unchanged using models addressing potential spatial autocorrelation. However, in two-pollutant models for the period 1995-2002, the associations between ozone and both cardiopulmonary mortality and AMI incidence were reduced to nonsignificance when either CO or NO_2 was in the model, or had a hazard ratio suggesting a significantly reduced risk when PM2.5 was in the model (Table 15). However, it should be noted that, due to the restrictions placed on spatial interpolations for both CO and NO_2 (Table 1), there were small numbers of participants (about 10 – 15% of the cohort) in all models involving these pollutants, including the two-pollutant models. Therefore, the two-pollutant models with either ozone and either CO or NO_2 excluded at least 80% of the participants and ozone observations, which suggests that these results should be interpreted with caution, at least with respect to the majority of the study participants who were excluded from these models.

In contrast, the two-pollutant models with ozone and PM2.5 included more than 90% of the participants. In these models (Tables 15 – 17), the PM2.5 hazard ratios retained their significance and either changed little or increased slightly, whereas those for ozone did not, regardless of whether imputed PM2.5 data were in the model (Tables 15 and 17) or not (Table 16). Both ozone and PM2.5 are regional pollutants and the individual long-term average exposures for these pollutants were strongly correlated ($r = 0.41$, $p < 0.0001$ for the period 1995-2002), suggesting that the ozone results may have been confounded by PM2.5. That similar results occurred in two-pollutant models restricted to the period 1999-2002 indicates that this phenomenon is not an artifact of imputation from the reconstruction of historical PM2.5 data. For cardiopulmonary mortality from 1999 through 2002, the ozone hazard ratio in the two-pollutant model with PM2.5 was 0.987 (95% CI = 0.979, 0.994), suggesting the possibility of multicollinearity having affected these results. We did not have the resources to fully explore the implications of the PM2.5-ozone relationship in this dataset, but recommend that additional analyses be undertaken in future research.

As noted above, the population sample sizes and numbers of events for the other gaseous pollutants were quite restricted relative to the analyses for ozone and PM due to the small radial distances we imposed on spatial interpolation (3 km for neighborhood-scale monitors and 5 km for urban/regional scale monitors (Table 1)). This decision was made in order to reduce exposure misclassification for these gaseous pollutants, which are subject to considerable intra-urban variability, depending largely on local traffic patterns. We recognize that in some instances even these relatively small buffer zones might be inadequate, given that NO_2 levels may sometimes vary several fold over a distance of several hundred meters (Singer et al. 2004). Despite (or maybe because of) this trade-off between statistical power and exposure misclassification, we identified

several associations between these traffic-associated gases and both all-cause mortality (NO_2 and CO in the full cohort for both exposure periods), and circulatory disease morbidity (NO_2 and CO for AMI and stroke incidence for both exposure periods, and NOx for AMI incidence for the longer exposure period). In two-pollutant models for AMI incidence, the CO hazard ratios markedly increased, regardless of the other pollutant in the model (PM2.5, ozone, or NO_2): the other pollutants' hazard ratios all decreased and became nonsignificant. In two-pollutant models for AMI, the NO_2 hazard ratios decreased, but nonetheless remained significant in the model with ozone. Considering how highly correlated the long-term average exposures to these pollutants were (except CO and ozone, Table 5), as well as the relatively small numbers of events in these two-pollutant models, the hazard ratios estimated from such models are more informative from a qualitative than a quantitative standpoint. In other words, it is likely that long-term exposures to CO and NO_2, particularly the former, are both associated with incidence of AMI, and probably with stroke (though we did not run two-pollutant models for the latter outcome). As the hazard ratios were relatively unaffected by the use of the full 1988-2002 versus the 1995-2002 exposure period, the additional years of estimated exposure are not likely to have been influential in these relationships. However, a more formal analysis of critical exposure windows would be useful in identifying the most relevant period(s) of exposure.

Both NO_2 and CO have been associated with myocardial infarction and stroke in time-series studies (Zanobetti and Schwartz 2006; Wellenius et al. 2005; Linn et al. 2000; Barnett et al. 2006). Coupled with concurrent associations of these outcomes with particulate matter, these results have generally been taken to signify that fossil fuel combustion represents the likely common source of all of these pollutants. In previously published cohort mortality studies, associations of these pollutants with mortality have been inconsistent. In the ACS-CPS II, neither pollutant was associated with all-cause or cardiopulmonary mortality. In the Dutch cohort study, effects of CO were not examined, but NO_2 was associated with cardiopulmonary mortality (Hoek et al. 2002). In the recent AHSMOG investigation, long-term NO_2 concentrations were associated with fatal CHD only in post-menopausal women; CO was likewise not examined in that publication (Chen et al. 2005).

Given our findings for traffic-related pollutants (PM2.5, NO_2, and CO), it was somewhat surprising that our analysis for several of the traffic metrics (distance to highway, traffic density at 150 and 300 meters) showed no association with any of the outcomes. However, it is possible that our approach of evaluating these metrics over their interquartile ranges may be partly responsible for the lack of association. Traffic-related responses are likely to be nonlinear: other studies have reported effects among those who resided in very close proximity to major roads (Hoek et al. 2002; Finkelstein et al. 2004). However, we did observe associations of road density with both all-cause mortality and stroke incidence, while the results for vehicle density suggested a weaker association with stroke incidence (HR = 1.05; 95% CI = 0.99-1.11). It should be noted that all of these basic traffic metrics are prone to error due to the counting procedures involved (e.g., within a 300 m buffer, most of the count might be close to the radial distance, with few vehicles actually passing close to a given residence). Moreover, recent cross-sectional traffic-related data were used to characterize the entire follow-up period. Therefore, unlike the pollutant distributions, which are functions of emissions (which

change over time) and other meteorological patterns and topography (which are relatively stable over the study period), traffic metrics may contain considerable error due to changes in road networks and traffic patterns. That is, the spatial patterns of local traffic metrics could change more than the measured pollutant concentrations, suggesting the possibility of greater misclassification error among the former.

We note two additional limitations to our results based on the use of interpolated pollutant surfaces. As described above and in Appendix 1, all monitors with sufficiently complete data were used to generate the monthly pollutant surfaces. This approach, while maximizing the spatial coverage in relation to the subjects' addresses, resulted in a dataset based on variable numbers of monitors over space and time, as they were deployed or taken out of operation, either transiently or permanently. Interpolation based on inverse distance weighting (as was done here) is quite sensitive to the underlying structure of the attribute data. In other words, the addition of a new monitor can drive much of the local surface estimation, causing a "pocket" effect around the monitor (potentially giving disproportionate weight to the information generated by that monitor compared to the pollutant surface estimation in the absence of the monitor), while dropping a monitor could potentially result in a "gapping" effect (in which more distant monitors would be disproportionately weighted). While we do not consider this likely to have had a marked effects on our results, we would recommend additional sensitivity analysis to assess the extent to which the use of different sets of monitors over time might have affected the generation of the pollutant surfaces.

Finally, our results, in general, did not appear to be affected by spatial autocorrelation, but we nevertheless experienced a few difficulties in assessing this phenomenon, especially in relation to reconstructed PM2.5 data. Examination of spatial autocorrelation in our data was intended to be exploratory: the Cox-Poisson random effects model that we used represents the state of the art in this field, but additional work is needed to refine its utility in the CTS and other datasets. Some issues to address include: visualization and assessment of both the exposure and residual risk surfaces, exclusion of areas with sparse event data, investigation of different autocorrelation matrices (e.g., based on distance rather than adjacency), and alternative characterization of the contextual variables (i.e., derived from Census data) at the zip code level.

Summary and Conclusions

In an ongoing cohort study of over 100,000 female participants in the California Teachers Study (CTS), we developed estimates of long-term air pollution exposure at the subjects' residences and examined associations between these exposure estimates and the following health outcomes: total mortality, cardiopulmonary mortality, incidence of acute myocardial infarction (AMI) and stroke. In addition, we examined the potential impacts of several traffic metrics on these outcomes. In order to derive the pollutant exposure metrics, the CTS participants' addresses were linked with monthly estimates of long-term exposure to multiple air pollutants, including PM10, PM2.5, and several gases (ozone, carbon monoxide, nitrogen oxides, nitrogen dioxide, and sulfur dioxide), as well as with several cross-sectional measures of traffic-related exposures from the year 2000 or later. The monthly pollutant surface estimates were provided by ARB staff. PM2.5 data before 1999 were derived under a separate contract managed by ARB. These pre-1999 data

were estimated using regression methods based on statistical associations of PM2.5 with other measures of particulate matter and with other pollutants. The main pollutant exposure periods used in this analysis were 1995 through 2002 (encompassing only the period after the CTS cohort was established) and 1988 through 2002. Residential addresses were not available for study participants prior to 1995.

The statistical analysis was conducted using Cox proportional hazard regression models, adjusting for covariates that have been used in other air pollution cohort studies and that have been found to be important in previous studies of the CTS, including: smoking status, total pack-years (for current and former smokers), body mass index, marital status, alcohol consumption, second-hand smoke exposure at home, dietary fat, fiber and calories, physical activity, menopausal status, hormone therapy use, family history of myocardial infarction and stroke, use of blood pressure medication, aspirin use, and several Census-derived contextual (neighborhood) variables (income, income inequality, education, population size, racial composition, unemployment). The analysis included estimation of hazard ratios for the whole cohort, and separately for those who did not move during the follow-up period (1995-2002) ("nonmovers") and for movers. We also undertook several sensitivity analyses, including: (i) examining the data for evidence of spatial autocorrelation; (ii) limiting the PM2.5 analysis to measured values only (1999-2002), eliminating estimated PM2.5 data from prior years; (iii) running two-pollutant models for cardiopulmonary mortality and AMI incidence for the major pollutants; (iv) examining two-pollutant models for PM2.5 and ozone, with both limited only to the period 1999-2002; and (v) examining the impact of using values of ozone measured only during the third quarter average (summer), when people are likely to spend more time outdoors and to have their windows open.

The most notable results from these analyses include the following:

1. PM2.5 was associated with all-cause and cardiopulmonary mortality, as well as incidence of AMI and stroke. The estimates for cardiopulmonary mortality and AMI incidence remained elevated and statistically significant in two-pollutant models. However, these estimates were based in part on historical reconstruction of PM2.5 levels prior to 1999, and were of considerably greater magnitude than those in all other published studies except Chen et al. (2005).

2. The method of developing the historical PM2.5 database may have led to systematic underestimation of the variance of actual PM2.5 concentrations and therefore overestimation of the associated hazard ratios. Thus, without further quantitative investigation of the extent of the measurement error introduced during the creation of this database, we believe that the use of these estimated values of PM2.5 in epidemiological investigations should be limited and that any results based on the use of these data should be interpreted with caution. Therefore, the analysis using historically reconstructed values of PM2.5 provides qualitative evidence of an association with all four health outcomes, but should not be construed as provided defensible quantitative estimates at this time.

3. In single-pollutant models using only measured data, PM2.5 was still associated with these four outcomes, with the following hazard ratios for a 10 $\mu g/m^3$ increase in long-term exposure: all-cause mortality 1.08 (95% CI 1.00 – 1.16); cardiopulmonary mortality 1.13 (95% CI 1.02 – 1.25); AMI incidence 1.14 (95% CI 1.02 – 1.27); stroke incidence 1.14 (1.00 – 1.29). These estimates were modestly increased by the inclusion of ozone in the models (for cardiopulmonary mortality and AMI) and are consistent with results from several previous studies of the effects of long-term exposure to PM2.5.

4. There were no significant associations between PM10 and any outcomes when the exposure period was limited to 1995 through 2002. In contrast, with the longer exposure period (1988-2002), PM10 exposures were significantly associated with all four outcomes. These PM10 hazard ratios were of similar magnitude to those for the measured PM2.5 data for the period 1999-2002, and were much lower than the hazard ratios involving the historical PM2.5 data. The estimated risk of death during the follow-up period per 10 $\mu g/m^3$ PM10 (1.09, 95% CI 1.04-1.13) is comparable to that reported for PM10 in the ACS CPS-II study (Pope et al. 2002).

5. Single-pollutant models for ozone indicated an association with AMI incidence with the 1995-2002 exposure period, and with all four outcomes when the longer exposure period (1988-2002) was used. However, with PM2.5, NO_2, or CO in the model, ozone was no longer positively associated with the adverse outcomes.

6. Due to the restrictions placed on spatial interpolations for both CO and NO_2, there were small numbers of participants (about 10 – 15% of the cohort) in all models involving these pollutants. Nevertheless, these traffic-associated gases were associated with all-cause, but not cardiopulmonary, mortality (for both exposure periods), and with circulatory events (NO_2 and CO with AMI and stroke incidence for both exposure periods, and NOx with AMI incidence for the longer exposure period). Even though the CO hazard ratios remained elevated and significant in two-pollutant models, as did NO_2 with ozone in the model, strong correlations involving these and other pollutants suggest that these results be interpreted with caution. Sulfur dioxide was not associated with any adverse outcomes except for cardiopulmonary mortality among nonmovers during the 1995-2002 exposure period.

7. In light of the associations between several traffic-related pollutant exposures (PM2.5, NO_2, and CO) and mortality and morbidity, it was somewhat surprising that our analysis for several of the traffic metrics (distance to highway, traffic density at 150 and 300 meters) showed no association with any of the outcomes. However, it is possible that the lack of an association was due to our having evaluated these metrics as continuous variables within buffers of 150 and 300 meters around each residence. More associations might have been detected had we considered smaller buffers (e.g., 100 m from major highways) and also tested for nonlinearity by modeling the effects of the extremes of the traffic density distributions (i.e., the top 10%). In addition, several characteristics of such traffic metrics may contribute to substantial exposure measurement error, reducing the likelihood of being able to detect associations. However, we did observe associations of

road density with both all-cause mortality and stroke incidence, while the results for vehicle density suggested a weaker association with stroke incidence.

8. There were greater proportions of movers who were younger, pre/peri-menopausal, and therefore less likely to have used hormone therapy, than among nonmovers. Nonmovers engaged in slightly less physical activity than movers, and were more likely to be married or living with a partner, and to report a family history of AMI or stroke. The regression analyses in general revealed relatively few differences in the pollutant-associated results between movers and nonmovers, except that the associations based on measured PM2.5 (1999-2002 only) with all four outcomes were of substantially greater magnitude among the subcohort of movers compared with nonmovers.

9. In general, the results of this analysis did not appear to be affected by spatial autocorrelation, but we nevertheless experienced difficulties in assessing this phenomenon for PM2.5 in particular. While the Cox-Poisson random effects model that we used represents the state of the art in this field, additional work is needed to refine its utility in the CTS and other datasets.

10. This investigation has both strengths and limitations.

a. Strengths of this analysis include:

(i) the large size of this cohort. Except for the ACS-CPS II, this is the largest air pollution cohort study undertaken to date.

(ii) the low prevalence of active smoking among the study participants.

(iii) the large proportion of women at risk of developing cardiovascular disease by virtue of their age and post-menopausal status. Other recent studies suggest that post-menopausal women appear to represent a particularly susceptible subgroup with respect to air pollution and cardiovascular disease.

(iv) the relative uniformity of occupational status. Therefore, the need to control statistically for the kinds of potentially toxic exposures that would be common in industrial environments was unnecessary. By comparison, the prevalence of inhalation occupational hazards ranged from 28 – 53% in the Harvard Six Cities study (Dockery et al. 1993).

(v) the unparalleled temporal and spatial resolution of pollutant exposures. No previous study has developed monthly exposure averages at the study participants' residential addresses.

(vi) the ability to examine incidence of AMI and stroke, not just fatal events, via linkage with comprehensive hospitalization as well as mortality data in California.

b. Limitations of this study include:

(i) probable overestimation of PM2.5-associated hazard ratios when historical reconstructions rather than measured PM2.5 data were used.

(ii) limitation of the study population to one gender only.

(iii) unknown error introduced into the development of pollutant surfaces by the use of all available monitors for each pollutant for the inverse distance weighted interpolation. While maximizing the spatial coverage in relation to the subjects' addresses, this approach resulted in a dataset based on variable numbers of monitors over time and space, as they were deployed or taken out of operation. As interpolation based

on inverse distance weighting is quite sensitive to the underlying structure of the attribute data, additional sensitivity analyses are necessary to understand whether this process may have affected our results.

(iv) use of cross-sectional traffic metric data from 2000 and later to estimate traffic exposures throughout the entire follow-up period. In addition, a more nonlinear analysis of the traffic data might provide greater insight into possible relationships with the health outcomes.

Recommendations

Based on the results of this investigation we provide the following recommendations for additional research involving the CTS cohort, as well as for the exposure data used in this investigation:

1. The use of substantial numbers of imputed values based on regression models to estimate chronic exposures produced risk estimates substantially greater than in most other studies of fine particles. This may be due to underestimation of the variance of real exposures, which in turn will result in overestimation of PM2.5-associated effects. We recommend that ARB undertake additional analyses of the historical PM2.5 database, using both simulations and real data, to gain a better understanding of the magnitude of the potential exposure measurement error introduced through the use of regression-based imputed fine particle data.
2. Identifying critical windows of exposure associated with mortality and cardiovascular morbidity would be important in estimating the benefits of controlling specific pollutants. We did not have the resources in this investigation to undertake such analyses, but recommend that additional efforts be undertaken to explore the magnitude of hazard ratio differences between different exposure periods prior to death or incident AMI or stroke. To the extent that there may be clear differences between the different exposure periods, this could provide data of high utility in benefits analysis.
3. In order to rule out effect modification by a history of active smoking, we recommend additional sensitivity analyses related to smoking status, including a set of analyses involving never-smokers only, or approximately two-thirds of this study population. Such analyses could also involve more refined variables representing exposure to second-hand smoke.
4. In this dataset, ozone was strongly associated with all adverse outcomes for the full exposure period in single-pollutant models, as well as with AMI incidence in the period from 1995-2002. However, these results appear to have been confounded by PM2.5 and possibly to have been subject to the effects of multicollinearity between these pollutants. We were unable to fully investigate the nature and extent of the how the relationships between these two pollutants might have affected their associations (or lack thereof) with the disease outcomes under study. In order to reduce the uncertainty about the influences of each of these pollutants on circulatory disease, we recommend that additional analyses be undertaken to examine whether there is potentially any effect of ozone on, e.g., AMI, independent of its correlation with PM2.5.
5. As the changing composition of the sets of monitors used in the IDW interpolation may have introduced an unknown error component into the pollutant surface estimation, we recommend additional sensitivity analysis to examine how the variability in monitor number might have affected the pollutant surfaces and the associated risk estimates. This would involve, e.g., identifying a core set of monitors in operation for at least 75% of the follow-up period, producing the pollutant surfaces from this core set of monitors, and comparing them with those produced using the variable numbers of monitors available for each pollutant.

6. To improve the utility of the Cox-Poisson random effects model used to assess spatial autocorrelation in our data, we recommend additional sensitivity and other analyses mentioned in the text, including: expanded visualization and assessment of both the exposure and residual risk surfaces, formal testing for local spatial autocorrelation in the residual relative risks, exclusion of areas with sparse event data to avoid convergence problems experienced in this analysis, investigation of different autocorrelation matrices (based on distance rather than adjacency), and additional characterization of the contextual variables (i.e., derived from Census data) at the zip code level, which matches the cluster level of the random effects model. As the contextual covariates were formulated at a different (Census tract) scale than the zip code clusters, this might have resulted in non-independence of the residuals within the zip code clusters, which violates the assumption of independent observations underlying statistical inference. There is no direct evidence of this problem in the results, but further investigation is warranted to ensure efficient hypothesis tests.

7. In this investigation we examined broad mortality categories only. We recommend additional analyses examining specific subclasses of causes of death (e.g., ischemic heart disease and stroke) to compare with other recent studies and to provide insight into pollutant-associated mechanisms of toxicity.

8. Since PM2.5 is a heterogeneous mix of particle sizes and chemistry, the specific constituents of concern have not been elucidated. This is a critical issue that may help: (1) target pollution control efforts and thereby reduce costs of control; (2) inform the work on biological mechanisms and plausibility; and (3) make it possible to conduct more targeted analysis of the health and economic benefits of controlling specific constituents of PM2.5 such as diesel particles, nitrates and metals. A number of California counties have PM2.5 species data going back to the year 2000. Therefore, we recommend that a dataset of long-term exposures to PM constituents be created and utilized in subsequent analyses of the CTS cohort.

9. In several other studies, the effects of traffic exposure metrics have tended to be quite nonlinear. We recommend additional examination of the traffic measures used in this study, using alternative buffers that emphasize the likelihood of very high exposures, such as residence within 100 meters of a highway.

10. One of the important assumptions underlying our analytical model is that the effect of each continuous predictor on the (logarithm of the) hazard is assumed to be linear. Misspecification of this assumption may result in a poorer model fit and the possibility of residual confounding. In addition, it is of particular interest to determine the actual shape of the concentration-response function to see if it departs significantly from linearity. Therefore, it is important to explicitly examine the linearity assumptions underlying the original analysis using a flexible spline regression model, which we would apply to PM2.5 and possibly other pollutants.

11. An additional critical question is what individual characteristics put individuals at risk from exposure to air pollution. Therefore, we recommend addressing this question by stratifying the teachers cohort, based on some of the results, into

certain subgroups (defined, for example, by their BMI, use of hormone therapy, whether they moved during the study period, and so forth) and examine whether the pollutant effects differ by subgroup. This analysis of effect modification may elucidate potential biological mechanisms underlying the effect estimates, aid in estimating the benefits of pollutant control, and help focus future research efforts.

References

Abbey DE, Nishino N, McDonnell WF, Burchette RJ, Knutsen SF, Beeson WL, Yang JX. Long-term inhalable particles and other air pollutants related to mortality in nonsmokers. Am J Respir Crit Care Med 1999; 159:373-382.

Alexandersen P, Tanko LB, Bagger YZ, Qin G, Christiansen C. The long-term impact of 2-3 years of hormone replacement therapy on cardiovascular mortality and atherosclerosis in healthy women. Climacteric. 2006;9:108-18.

Analitis A, Katsouyanni K, Dimakopoulou K, Samoli E, Nikoloulopoulos AK, Petasakis Y, Touloumi G, Schwartz J, Anderson HR, Cambra K, Forastiere F, Zmirou D, Vonk JM, Clancy L, Kriz B, Bobvos J, Pekkanen J. Short-term effects of ambient particles on cardiovascular and respiratory mortality. Epidemiology. 2006;17:230-3.

ARB/OEHHA. California Air Resources Board and Office of Environmental Health Hazard Assessment. Review of the California Ambient Air Quality Standard For Ozone, Sacramento, CA: March 11, 2005.

Barnett AG, Williams GM, Schwartz J, Best TL, Neller AH, Petroeschevsky AL, Simpson RW. The effects of air pollution on hospitalizations for cardiovascular disease in elderly people in Australian and New Zealand cities. Environ Health Perspect 2006;114:1018-23.

Bell ML, Samet JM, Dominici F. 2004. Time-series studies of particulate matter. Annu Rev Public Health. ;25:247-80.

Bernstein L, Allen M, Anton-Culver H, Deapen D, Horn-Ross PL, Peel D, et al. High breast cancer incidence rates among California teachers: results from the California Teachers Study (United States). Cancer Causes Control 2002;13:625-35.

Blanchard C and Tanenbaum S. Spatial and temporal characterization of fine particulate matter mass concentrations in California, 1980-2002. Final Report to the California Air Resources Board, Contract Number 03-350, December 2005. (http://www.arb.ca.gov/research/apr/past/03-350a.pdf)

Brook RD, Franklin B, Cascio W, Hong Y, Howard G, Lipsett M, Luepker R, Mittleman M, Samet J, Tager I. Air pollution and cardiovascular disease: a statement for healthcare professionals from the Expert Panel on Population and Prevention Science of the American Heart Association.. Circulation 2004:109:2655-71.

California Department of Health Services, C-Stats (County and Statewide Archive of Tobacco Statistics), http://www.cstats.info/, accessed May 1, 2007.

Centers for Disease Control. Cigarette smoking among adults – United States 2000. MMWR 2002;51:642-645.

Centers for Disease Control. State-specific prevalence of cigarette smoking -- United States, 1995. MMWR 1996;45:962-966.

Centers for Disease Control. State-specific prevalence of current cigarette smoking among adults, and policies and attitudes about secondhand smoke --- United States, 2000. MMWR 2001;50:1101-06.

Chen LH, Knutsen SF, Shavlik D, Beeson WL, Petersen F, Ghamsary M, Abbey D. The association between fatal coronary heart disease and ambient particulate air pollution. Are females at greater risk? Environ Health Perspect 2005;113:1723-29.

Dhalla NS, Temsah RM, Netticadan T. Role of oxidative stress in cardiovascular diseases. J Hypertens 2000; 18(6):655-73.

Dockery DW, Pope CA III, XU X, Spengler JD, Ware JH, Fay ME, Ferris BG, Speizer FE. An association between air pollution and mortality in six U.S. cities. N Engl J Med 1993;329:1753-1759.

Donaldson K, Stone V, Seaton A, MacNee W. Ambient particle inhalation and the cardiovascular system: potential mechanisms. Environ Health Perspect 2001; 109 (Suppl 4):523-27.

Enstrom JE. Fine particulate air pollution and total mortality among elderly Californians, 1973-2002. Inhal Toxicol 2005:17:803-816.

Finkelstein MM, Jerrett M, Sears MR. Traffic air pollution and mortality rate advancement periods. Am J Epidemiol 2004;160:173-7.

Garbe E, Suissa S. Hormone replacement therapy and acute coronary outcomes: methodological issues between randomized and observational studies. Hum Reprod 2004;19:8-13.

Gurgueira SA, Lawrence J, Coull B, Murthy GG, Gonzalez-Flecha B. Rapid increases in the steady-state concentration of reactive oxygen species in the lungs and heart after particulate air pollution inhalation. Environ Health Perspect 2002;110:749-55.

Health Effects Institute. Reanalysis of the Harvard Six Cities study and the American Cancer Society study of particulate air pollution and mortality: A special report of the Institute's Particle Epidemiology Reanalysis Project. Cambridge, MA: 2000, 293 pp.

Hedblad B, Merlo J, Manjer J, Engstrom G, Berglund G, Janzon L. Incidence of cardiovascular disease, cancer and death in postmenopausal women affirming use of hormone replacement therapy. Scand J Public Health 2002;30:12-9.

Hoek G, Brunekreef B, Goldbohm S, Fischer P, van den Brandt PA. Association between mortality and indicators of traffic-related air pollution in the Netherlands: a cohort study. Lancet 2002; 360:1203-9.

Jaro MA. Probabilistic linkage of large public health data files. Stat Med 1995; 14:491-498.

Jerrett M, Burnett RT, Ma R, Pope CA, Krewski D, Newbold KB, et al. Spatial Analysis of Air Pollution and Mortality in Los Angeles. Epidemiology 2005;16:1-10.

Krewski D, Burnett RT, Goldberg MS, Hoover BK, Siemiatycki J, Jerrett M, Abrahamowicz M, White WH. Overview of the reanalysis of the Harvard Six Cities Study and American Cancer Society Study of Particulate Air Pollution and Mortality. J Toxicol Environ Health A. 2003;66:1507-51.

Kunzli N, Jerrett M, Mack WJ, Beckerman B, LaBree L, Gilliland F, Thomas D, Peters J, Hodis HN. Ambient air pollution and atherosclerosis in Los Angeles. Environ Health Perspect. 2005;113:201-6

Laden F, Schwartz J, Speizer FE, Dockery DW. Reduction in fine particulate air pollution and mortality. Am J Respir Crit Care Med 2006; 173:667-672.

Li N, Sioutas C, Cho A, Schmitz D, Misra C, Sempf J, Wang M, Oberley T, Froines J, Nel A. Ultrafine particulate pollutants induce oxidative stress and mitochondrial damage. Environ Health Perspect. 2003;111:455-60

Linn WS, Szlachcic Y, Gong H Jr, Kinney PL, Berhane KT. Air pollution and daily hospital admissions in metropolitan Los Angeles. Environ Health Perspect. 2000;108:427-34.

Lipsett M. Air pollution and cardiovascular disease in the California Teachers Study cohort. Proposal submitted to California Air Resources Board, Research Division, Sacramento, CA, January 12, 2004.

Ma R, Krewski D, Burnett RT. Random effects Cox models: a Poisson modelling approach. Biometrika 2003;90:157-169.

Nafstad P, Haheim LL, Wisloff 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;112:610-5.

OEHHA/ARB. Proposed identification of environmental tobacco smoke as a toxic air contaminant. Part B. health effects. As approved by the Scientific Review Panel, June 24, 2005. http://www.arb.ca.gov/toxics/id/summary/etspt_b.pdf

Pope CA III, 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. JAMA 2002; 287:1132-1141.

Pope CA 3rd, Burnett RT, Thurston GD, Thun MJ, Calle EE, Krewski D, Godleski JJ. Cardiovascular mortality and long-term exposure to particulate air pollution: epidemiological evidence of general pathophysiological pathways of disease. Circulation. 2004;109:71-7. Epub 2003 Dec 15.

Pope CA III, Thun MJ, Namboodiri MM, Dockery DW, Evans JS, Speizer FE, Heath CW. 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.

R Development Core Team. R: A language and environment for statistical computing. 2006. Vienna: R Foundation for Statistical Computing. http://www.R-project.org.

SAS Institute, Inc. 2000. SAS Version 9. Cary, NC, USA:SAS Institute, Inc.

Singer BC, Hodgson AT, Hotchi T, Kim JJ. Passive measurement of nitrogen oxides to assess traffic-related pollutant exposure for the East Bay Children's Respiratory Health Study. Atmos Environ 2004;38:393-403.

Sun Q, Wang A, Jin X, Natanzon A, Duquaine D, Brook RD, Aguinaldo JG, Fayad ZA, Fuster V, Lippmann M, Chen LC, Rajagopalan S. Long-term air pollution exposure and acceleration of atherosclerosis and vascular inflammation in an animal model. JAMA 2005;294:3003-10.

U.S. Environmental Protection Agency (U.S. EPA). 2004. Air quality criteria for particulate matter. Washington, D.C., EPA 600/P-99/002aF-bF.

von Klot S, Peters A, Aalto P, Bellander T, Berglind N, D'Ippoliti D, Elosua R, Hormann A, Kulmala M, Lanki T, Lowel H, Pekkanen J, Picciotto S, Sunyer J, Forastiere F. Health Effects of Particles on Susceptible Subpopulations (HEAPSS) Study Group. Ambient air pollution is associated with increased risk of hospital cardiac readmissions of myocardial infarction survivors in five European cities. Circulation 2005;112:3073-9. Erratum in: Circulation 2006;113(5):e71.

Waters DD, Alderman EL, Hsia J, Howard BV, Cobb FR, Rogers WJ, Ouyang P, Thompson P, Tardif JC, Higginson L, Bittner V, Steffes M, Gordon DJ, Proschan M, Younes N, Verter JI. Effects of hormone replacement therapy and antioxidant vitamin supplements on coronary atherosclerosis in postmenopausal women: a randomized controlled trial. JAMA. 2002;288:2432-40.

Wellenius GA, Schwartz J, Mittleman MA. Air pollution and hospital admissions for ischemic and hemorrhagic stroke among medicare beneficiaries.Stroke. 2005;36:2549-53. Epub 2005 Oct 27.

World Health Organization. WHO air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide. Global update 2005. Summary of risk assessment. http://www.who.int/phe/air/aqg2006execsum.pdf (2006)

Zanobetti A, Schwartz J. Air pollution and emergency admissions in Boston, MA. J Epidemiol Community Health. 2006;60:890-5.

Glossary of terms, abbreviations, and symbols

ACS	American Cancer Society
AHSMOG	Adventist Health and Smog Study
AMI	acute myocardial infarction
ARB	Air Resources Board
BMI	body mass index (weight/height^2)
CO	carbon monoxide
CHD	coronary heart disease
CP	cardiopulmonary
CPS	Cancer Prevention Study
CTS	California Teachers Study
FRM	Federal reference method
HPMS	Highway Performance Monitoring System
HR	hazard ratio
IQR	interquartile range
MI	myocardial infarction
NOx	nitrogen oxides
NO_2	nitrogen dioxide
O_3	ozone
OSHPD	Office of Statewide Health Planning and Development
PM	particulate matter
PM2.5	PM with a median aerodynamic diameter < 2.5 μ
PM10	PM with a median aerodynamic diameter < 10 μ
RR	relative risk
SO_2	sulfur dioxide
STRS	State Teachers Retirement System
USC	University of Southern California
USPS	United States Postal Service
VMT	vehicle miles traveled

Appendix 1 – Development of pollutant surfaces by ARB staff

The following pages contain brief descriptions and illustrative documentation of the process followed by ARB staff to create the monthly monitor averages and IDW pollutant surfaces.

METHODOLOGY

DATA

AIR POLLUTION DATA

•Obtained pollutant data from 1988 through 2002.

•Created monthly averages using complete day or representative day (See page titled: "Air Quality Data for Teacher's Cohort Study (CTS) Notes on Air Quality Data Statistics (1988-2002) June 15, 2004").

MONITOR DATA

•Obtained monitor site information primarily from the 2005 ARB's Web page NAMS/SLAMS document. In addition, made corrections to the latitude and longitude fields for some of the very old monitors based on their address when available.

JOINING AIR POLLUTION AND MONITOR DATA

•Air pollution and monitor data were linked using Microsoft Access

•Extracted only the data that met representative months based on the frequency of monitoring, and when the site scale was available, I did not include micro- or middle-scale sites. Used only neighborhood, regional, and urban scales. The final table was then exported to excel and saved in dbf format.

DEVELOPED INTERPOLATED RASTER SURFACES

•Used ARC Info., Spatial Analysis tool to create a model that was then exported into Python language.

•The interpolation was based on IDW using a fixed radius of 50,000 m and power of two for all pollutants.

•Created a Python script to calculate the IDW interpolation for every month between 1988 -2002.

•Final product consists of raster files with 250 m grids.

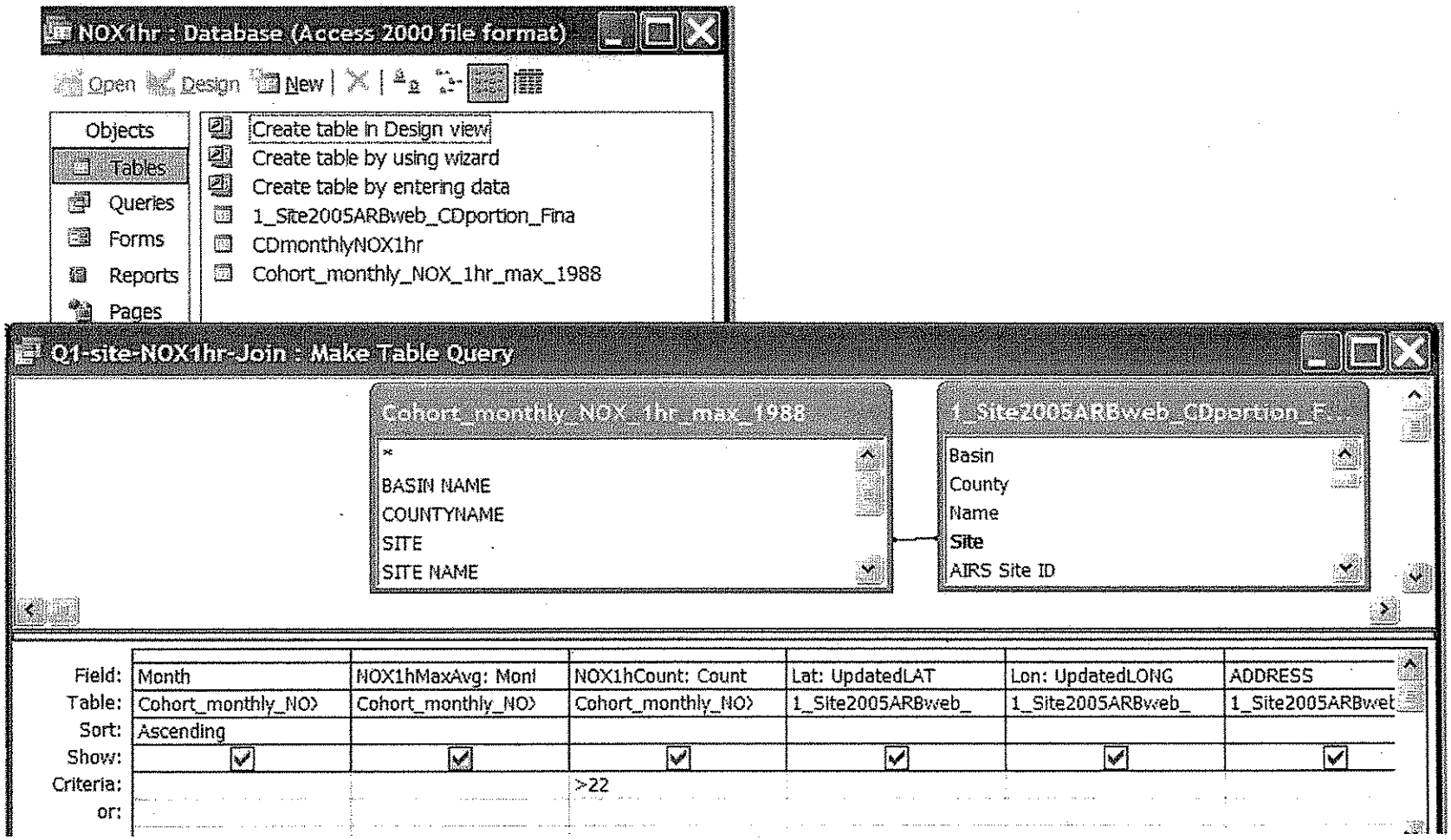
Air Quality Data for the Teacher's Cohort Study (CTS)
Notes on Air Quality Data Statistics (1988-2002) June 15, 2004

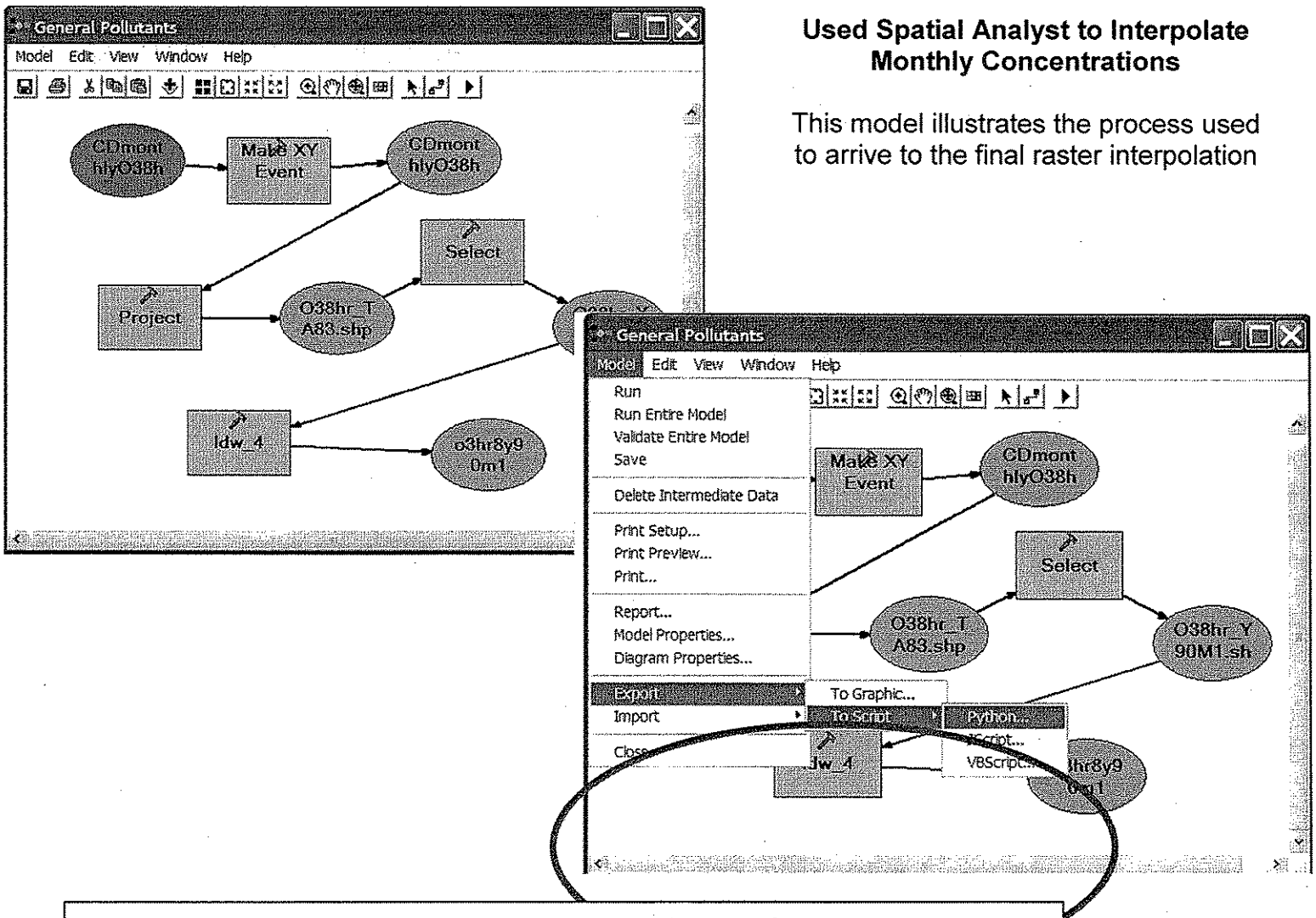
•PM10 mass data are based on 24-hour filter-based PM10 SSI (Size Selective Inlet) monitoring at standard conditions. The PM10 mass data are averaged across monitors for each site and date, and then the monthly averages are calculated based on the daily averages. PM10 SSI mass were usually monitored one in six days but the schedule may have varied from site to site.
•Ozone (1-hr. max.) – monthly averages are based on complete* days.
•NO2 (1-hr. max.) – monthly averages are based on complete* days.
•NOx (1-hr. max.) – most monthly averages are based on representative** days. The San Francisco Bay Area Air Basin sites had many exceptions. San Francisco Bay Area Air Basin sites 3659 and 3660 were based upon representative** days. The other San Francisco Bay Area Air Basin sites' NOx data for 1994-2002 were not reported, so the hourly data for NO and NO2 were summed for each hour to represent NOx. The daily maximum one-hour for NOx (San Francisco Bay Area Air Basin) included all the days available from the calculated NOx.
•NOx (24-hr. avg.) – monthly averages were based on representative** days. The 1994-2002 San Francisco Bay Area Air Basin NOx 24-hr. averages (except for sites 3659 and 3660) were derived from the summation of daily average NO2 and NO based upon representative** days.
•SO2 (24-hr. avg.) – monthly averages were based on complete* days. When the monthly average of SO2 is zero, the SO2 data for the month were below the limit of detection.
•CO (8-hr. avg.) – monthly averages were based on daily maximum eight-hr averages (non-overlapping State specification) for complete* days.
•* Complete Days: The monitoring satisfied state completeness criteria for area designation purposes for the specified pollutant for those days. http://www.arb.ca.gov/regact/areades/area00/attb.pdf
•** Representative Days: The data collected must be representative according to the following definition. There must be no more than two missing hours in any of the three consecutive eight-hour periods within a day. For an entire day, no more than two consecutive hours can be missed. Therefore, for an entire day, if there were three consecutive hours missed, the day would be invalidated.

NOX ONE HOUR MAX MONTHLY AVERAGE

Join Site Table w/ NOX Table

Included only months averages that had more than 22 monthly values





Used Spatial Analyst to Interpolate Monthly Concentrations

This model illustrates the process used to arrive to the final raster interpolation

The ArcMap Model was then converted to Python Script, an example of which can be found on the next page

```
gp.AddToolbox("C:/Program Files/ArcGIS/ArcToolbox/Toolboxes/Analysis Tools.tbx")

# Set the Geoprocessing environment...
gp.scratchWorkspace = "C:\\IDW\\pollutant\\NO2_1"
gp.outputCoordinateSystem = "PROJCS['TealeAlbersNAD83',GEOGCS['GCS_North_American_1983',DATUM['D_North_American_1983',SPHE
gp.outputZFlag = "Same As Input"
gp.clusterTolerance = ""
gp.extent = "-373976.780431 -604526.124980 540015.400156 450070.871569"
gp.outputZValue = ""
gp.outputMFlag = "Same As Input"
gp.workspace = "C:\\IDW\\pollutant\\NO2_1"
#********************************************************************************

# Local variables...
#NO21hrym_shp = "C:\\IDW\\pollutant\\NO2_1\\NO21hrym.shp"
NO21hrTA83_shp = "C:\\IDW\\pollutant\\NO2_1\\NO21hrTA83.shp"
NO21hr_Layer = "NO21hr_Layer"
CDmonthlyNO21hr_dbf = "C:\\IDW\\pollutant\\NO2_1\\CDmonthlyNO21hr.dbf"
#NO21hryrmo = "C:\\IDW\\pollutant\\NO2_1\\no21hryrmo"

# Process: Make XY Event Layer...
gp.MakeXYEventLayer_management(CDmonthlyNO21hr_dbf, "LON", "LAT", NO21hr_Layer, "GEOGCS['GCS_North_American_1983',DATUM['I

# Process: Project...
gp.Project_management(NO21hr_Layer, NO21hrTA83_shp, "PROJCS['TealeAlbersNAD83',GEOGCS['GCS_North_American_1983',DATUM['D_N

for year in range (1988, 2003):
    for month in range (1, 13):
        print year, month


        NO21hrym_shp = "C:\\IDW\\pollutant\\NO2_1\\NO21ym%d%d.shp" % (year,month)

        # Process: Select...
        gp.Select_analysis(NO21hrTA83_shp, NO21hrym_shp, "YEAR =%d AND MONTH =%d"% (year, month))

        NO21hrym = "C:\\IDW\\pollutant\\NO2_1\\NO21ym%d%d" % (year,month)

        # Process: IDW...
        gp.Idw_sa(NO21hrym_shp, "NO2HR1MAXA", NO21hrym, "250", "2", "FIXED,50000", "")
```

Appendix 2: Creation of Historical Pollutant Database by Dr. Charles Blanchard, (Excerpt from Blanchard and Tanenbaum 2005)

II. METHODS

Federal Reference Method (FRM) measurements of $PM_{2.5}$ mass concentrations (fine mass) are available beginning in 1998 or 1999. The US EPA has established a criterion for predictability of FRM fine mass concentrations from other measurements, which is a correlation coefficient of $r^2 \geq 0.8$. We use this criterion to select measurements suitable for prediction of FRM-equivalent fine mass concentrations. In general, mass concentration measurements, while meeting a criterion of predictability, need not be equivalent to FRM concentrations; they may exhibit either additive or multiplicative biases relative to FRM fine mass concentrations (Motallebi et al., 2003a; 2003b). We established conversion factors to standardize fine mass measurements from other networks to FRM equivalents. The other networks include the California Air Resources Board (CARB) dichotomous sampler network and a variety of special studies conducted prior to implementation of the FRM network (see Figures 1 through 5):

– CalTech – 1982, 1986, 1993; 5-11 sites, SoCAB
– IMPROVE – 1987 – 2002; 8-13 sites, state
– Valley Air Quality Study (VAQS) – 1988 – 1989; 6 sites, SJV
– California Acid Deposition Monitoring Program (CADMP) – 1988 – 1995: 10 sites and 1995-99, 5 sites, statewide
– Two-week sampler (TWS) – 1994 – 2002; 12 sites, SoCAB
– Integrated Monitoring Study, 1995 (IMS95) – 12/95 – 1/96; 10 sites, central CA
– PM Enhancement Program (PTEP) – 1995 – 1996; 6 sites, SoCAB

The principal constituents of $PM_{2.5}$ mass in California are organic and black (elemental) carbon, sulfate, and nitrate (McMurry et al., 2004). These PM components, in turn, are typically found primarily in the fine fraction. As a result, it is possible to reconstruct fine mass concentrations and their uncertainties at places and during times without measurements of $PM_{2.5}$ mass using measurements of sulfate, nitrate, and carbon from PM_{10} samples. CARB has developed a substantial monitoring record of PM_{10} sulfate and nitrate concentrations, but PM_{10} measurements of total carbon are limited to a few sites and years. We therefore established correlations between total carbon and related measurements, namely, coefficient of haze (CoH) and carbon monoxide (CO).
We also investigated the comparability of light extinction measurements (nephelometer data) and fine mass concentrations. Fine mass concentrations and nephelometer measurements were well correlated ($r^2 \geq 0.8$) during the years 1988 – 1994, but were poorly correlated ($r^2 \sim 0.4$) from 1995 - 2002.

In developing monthly averages of measured and reconstructed fine mass constructions, we established a selection priority as follows:

1. FRM fine mass
2. dichotomous sampler fine mass
3. CADMP fine mass and fine mass from other special studies
4. reconstruction from PM_{10} sulfate + nitrate + total carbon
5. reconstruction from PM_{10} sulfate + nitrate + total carbon calculated from CoH
6. reconstruction from PM_{10} sulfate + nitrate + total carbon calculated from CO
7. reconstruction from nephelometer data prior to 1995

For each day of a month, a daily-average PM level was obtained following the preceding priorities. Then, a monthly average was determined from all days in a month having data.

Appendix 3: ARCGIS Algorithms for Development of Several Exposure Metrics

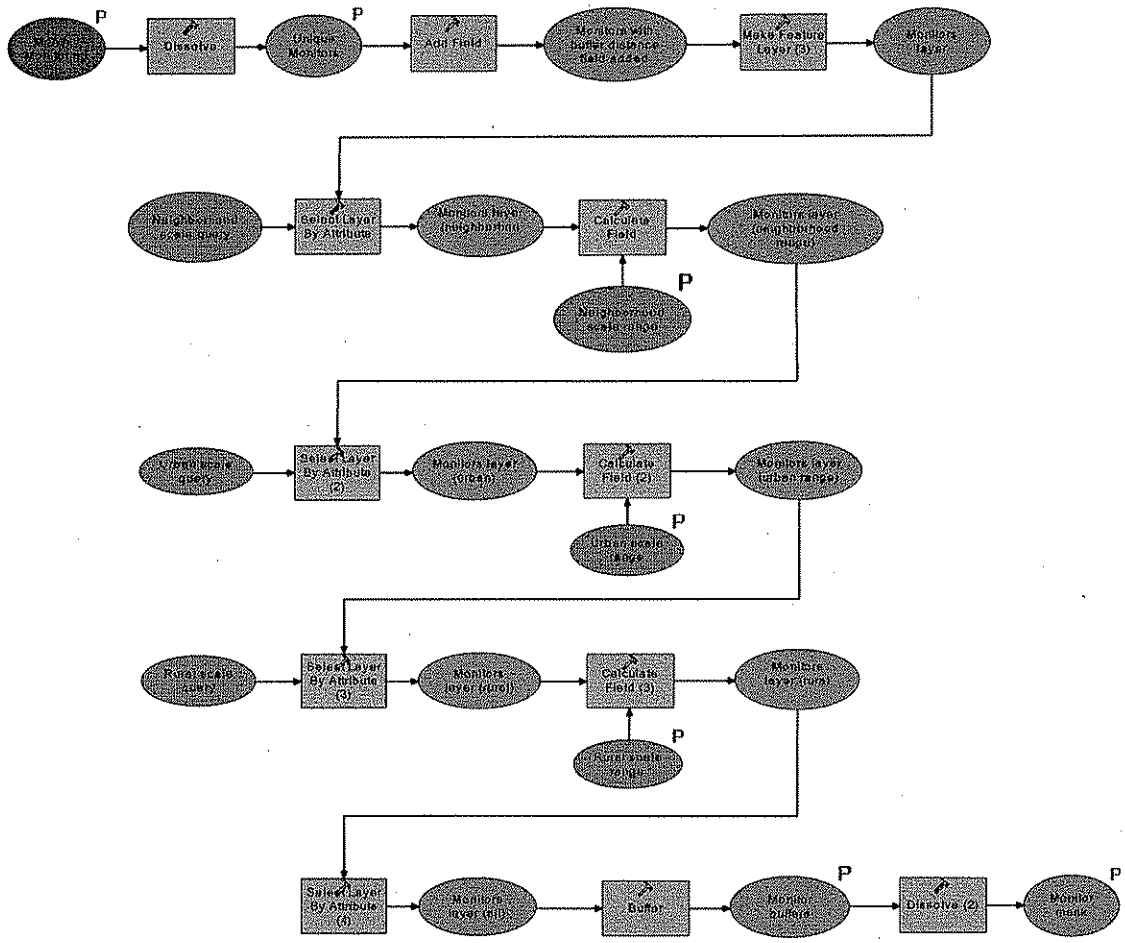


Figure 3-1: ArcGIS model for creating a mask (spatial filter) based on the spatial scale of each monitor.

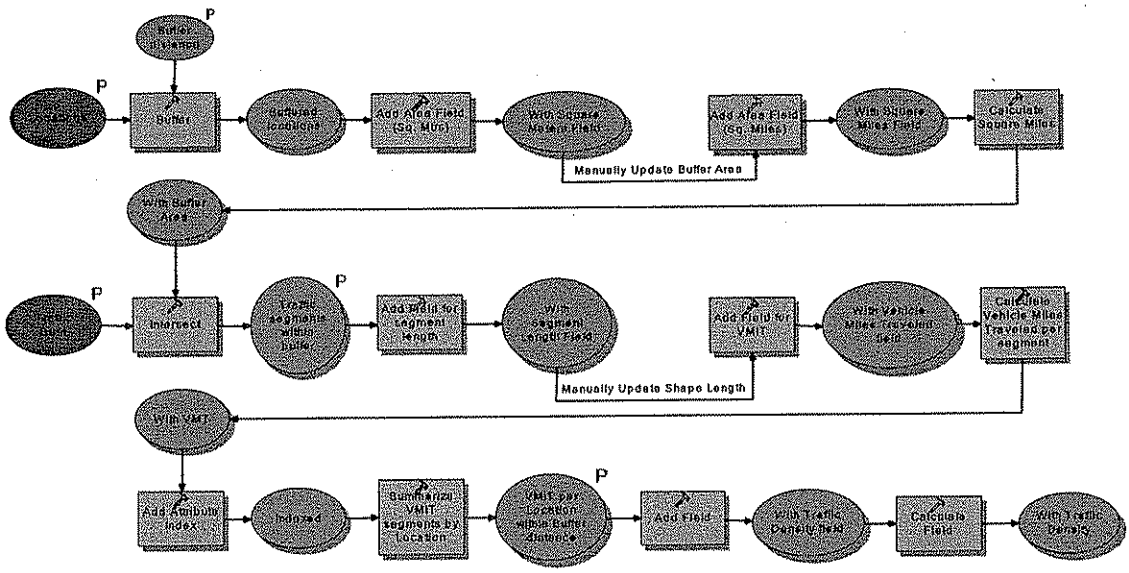


Figure 3-2: ArcGIS model for calculating traffic density

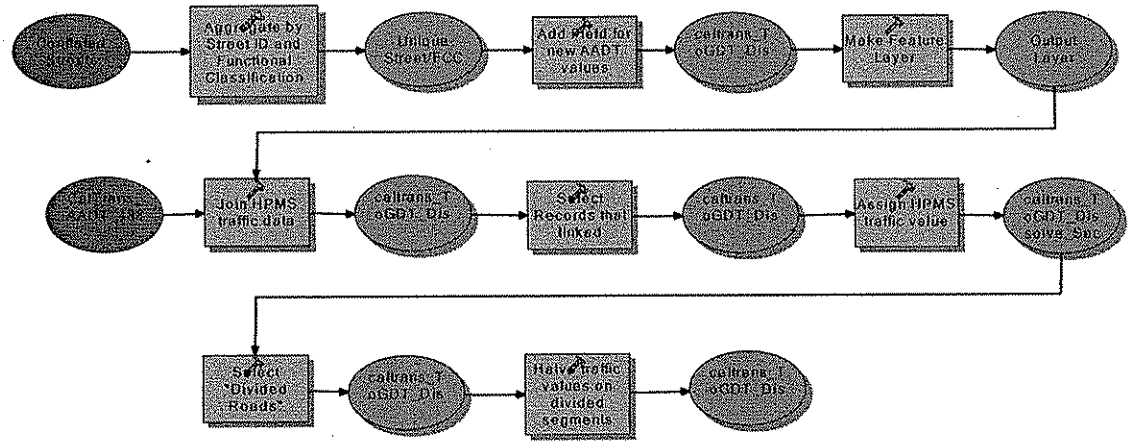


Figure 3-3: ArcGIS model for linking HPMS to conflated Dynamap streets