HEI

APPENDIX AVAILABLE ON THE HEI WEB SITE

Research Report 190

The Effects of Policy-Driven Air Quality Improvements on Children's Respiratory Health

Gilliland et al.

Appendix A: Ambient Air Quality Measurements and Supplemental Methods

This Appendix was reviewed solely for spelling, grammar, and
cross-references to the main text. It has not been formatted or fully edited by HEI.

This document was reviewed by the HEI Review Committee.

Correspondence may be addressed to Dr. Frank Gilliland, 2001 N. Soto Street, MC 9237, Los Angeles, CA 90089;
e-mail: *gillilan@usc.edu*.

APPENDIX A. Ambient Air Quality Measurements and Supplemental Methods

Frank Gilliland[1*], Edward Avol[1], Rob McConnell[1], Kiros Berhane[1], W. James Gauderman[1], Fred W. Lurmann[2], Robert Urman[1], Roger Chang[1], Edward B. Rappaport[1], Stephen Howland[1]

[1]University of Southern California, Dept. of Preventive Medicine, Los Angeles, CA
[2]Sonoma Technology Inc, Petaluma, CA

*Corresponding author: 2001 N. Soto Street, MC 9237, Los Angeles CA 90089, Tel: 323-442-1309, Fax: 323-442-3272, email: gillilan@usc.edu

CONTENTS

NETWORK OBJECTIVES AND OPERATIONS

The objective of the Children's Health Study (CHS) ambient air monitoring program was to obtain seasonal and annual average concentrations for air pollutants of CHS interest at a representative location in each CHS community. These data were then to be used for subsequent investigation of air pollution and chronic health effect associations. An additional objective was to obtain hourly and daily data for a subset of pollutants that could be used in assessment of potential chronic effects resulting from repeated short-term exposures. Continuous hourly measurements of ozone, NO_2, NO, and PM_{10} were made at the central CHS air monitoring station in each community. Integrated measurements of $PM_{2.5}$ mass, PM chemical constituents, and gaseous acids were made using a multi-legged two-week sampler (TWS) designed for the study (Lurmann et al. 1994). The PM chemical constituents included $PM_{2.5}$ sulfate, nitrate, and ammonium and PM_{10} elemental carbon (EC) and organic carbon (OC). The filters collected for carbon analysis were initially collected without a specified size-cut, and later using dual sampling legs with a $PM_{2.5}$ size cut and the undifferentiated leg (for comparative purposes). The TWS also collected samples for subsequent laboratory determination of two-week average concentrations of nitric acid, hydrochloric acid, formic acid, and acetic acid. These measurements were made throughout the study period, from 1994–2004. Additional measurements of carbon monoxide (CO), particle number (PN), $PM_{2.5}$ EC, $PM_{2.5}$ OC, and $PM_{2.5}$ elements by X-ray fluorescence (XRF) were implemented in most communities in 2000 and 2001. Hourly temperature and relative humidity were measured at some of the CHS air monitoring stations for some of

the years to complement the air quality data. These data were supplemented with meteorological data collected at locations (regional air monitoring sites, airports, and similar sources of credible meteorological data) near the CHS communities. The CHS air quality monitoring network was established by augmenting existing regional air monitoring stations and creating new stations in late 1993 and early 1994. Seven air monitoring stations were previously existing sites, where air pollution control agencies monitored ozone, NO_2, or PM_{10}. These existing sites included Atascadero, Santa Maria, North Long Beach, Lancaster, Upland, Lake Elsinore, and Alpine (see Figure A.1). Five new air monitoring sites were established for the study. These new sites were located at the U.C. Riverside Agricultural Station in Riverside, Jurupa Valley High School in Mira Loma, Rim of the World High School in Lake Arrowhead, Cabrillo High School in Lompoc, and Gladstone Elementary School in San Dimas. All the station locations met EPA siting requirements.

Figure A.1. Map of CHS communities.

Several stations were relocated, out of necessity, during the performance of the longitudinal study. The community air monitoring station at Gladstone Elementary School in San Dimas was shut down in 1996 at the request of the school district, on the basis of unspecified neighborhood complaints regarding operating noise. The CHS air monitoring instrumentation was moved to the nearest district-operated regulatory air monitoring station (in Glendora), which was approximately 4 km northwest of the original San Dimas sampling site. Prior to station shutdown and instrument relocation, samplers and instruments were operated concurrently at both locations for 12 months to characterize typical air quality differences between the San Dimas and Glendora locations. Based on the accrued co-operating sampling data, it was determined that

ambient ozone, NO_2, $PM_{2.5}$, and PM_{10} for San Dimas could be accurately estimated from Glendora data and the nearby Azusa station data (5 km west of San Dimas). However, due to localized differences in proximity to roadways, NO in San Dimas could not be accurately estimated from the other nearby stations.

Other station relocations did not always allow for one-year concurrent co-operation and data comparisons. In 2000, the California Air Resources Board (ARB) relocated the Santa Maria regional air monitoring station 0.4 km southwest of the original location. In 2001, the Antelope Valley Air Quality Management District moved the Lancaster station 2.3 km south of its original location. Concurrent sampling data were not collected when the Santa Maria and Lancaster stations were moved. However, both the new and old sampling locations in Santa Maria and Lancaster were similar with regard to roadway proximity, traffic density, spatial topology, and general land use (in terms of potential local sources), so differences in measured ambient air concentrations between the old and new sampling locations were likely to be small.

Additional changes were made to both the monitoring network and the study infrastructure as the study years passed by. When research funding transitioned from ARB sponsorship to NIEHS sponsorship in 2002, four additional study communities were added in Glendora, Anaheim, San Bernardino, and Santa Barbara. The routine air monitoring stations in these communities were augmented with additional equipment comparable to the other CHS stations. These four new communities replaced Lompoc, Lancaster, and Atascadero, whose school districts elected to no longer participate in the multi-year study. Following cessation of active study participants in these three

communities, air quality data from these sites were no longer archived in the study database after 2004.

The CHS air monitoring station at Rim of the World High School in Lake Arrowhead (5700 ft elevation) was shut down in 2004 and replaced with data collection activities from the Lake Gregory regional monitoring station, located 6 km west at a similar elevation (~4800 ft). Study-supported monitoring activities at the school-based station at Jurupa Valley High School in Mira Loma were taken over by the South Coast Air Quality Management District (SCAQMD) in 2004. Due to school growth and space limitations, the SCAQMD subsequently decommissioned the high school site in 2006 and established a new Mira Loma regional monitoring station at Van Buren Elementary School, 3 km to the southeast. In Riverside, after the special monitoring equipment was removed from the UCR Agricultural Station in late 2004, data collected at the extensively equipped regional trend station operating at Rubidoux, CA were used until the Magnolia station in central Riverside was augmented with suitable monitoring equipment.

Across the sampling network, there was also some evolution of sampling instrumentation over the two decades of study operations. After the CHS-developed two-week samplers were shut down due to lack of funding and an ARB request for instrumentation return, community monitoring stations were equipped with $PM_{2.5}$ beta attenuation monitoring samplers (BAMs) if the respective stations were not already equipped with $PM_{2.5}$ Federal Reference Method-approved (FRM) filter samplers. The PM_{10} tapered element oscillating microbalances (TEOMs), originally deployed in 1993 to provide continuous PM sampling data (as compared with the time-integrated approach

of the CHS two-week samplers), were also replaced with PM_{10} BAMs in 2009. Likewise, the NOx and ozone analyzers were replaced as needed over the course of the study.

Responsibility for field operations and data management evolved over the study period, as well. Initially, field operations were divided among Sonoma Technology, Inc. (STI), CARB, SCAQMD, San Luis Obispo County Air Pollution Control District's (SLOAPCD), and San Diego County Air Pollution Control Agency (SDCAPCD). In 1995, CARB contracted with the local air pollution agencies and RM Environmental Inc. for field operations of the stations through September 2004. The air quality data flowed from the field operations group to the exposure data manager, which was STI in 1993–1994 and ARB in 1995–2004. After 2004, the field operations and data management were primarily conducted by local districts and the CARB with help from STI. Final quality assurance of all exposure data was provided by STI through 2012.

Filters and coated-denuders were prepared and processed by the Environmental Health Service at the Los Amigos Research and Education Institute (LAREI). Gravimetric analyses of $PM_{2.5}$ mass and ion chromatography for determination of sulfate, nitrate, ammonium, formate, and acetate were conducted in the LAREI laboratory. The thermal/optical transmittance (TOT) laboratory analyses of EC and OC were performed at the California Institute of Technology. The X-ray Fluorescence (XRF) elemental analysis of $PM_{2.5}$ was conducted at the Desert Research Institute (DRI). The exposure data manager delivered all the data to STI for quality assurance and data validation prior to delivery of the data to the CHS research team, which was led by University of Southern California (USC) investigators. For data acceptance criteria, the U.S. EPA's 75% completeness criteria at the daily and monthly levels were applied.

Eight-hour average daily maximum ozone concentration data were required to have at least six hours of valid data (out of eight consecutive daytime hours) for acceptance. Twenty-four-hour average concentration data were required to have at least 18 hours of valid data for acceptance. Monthly-average data was required to have at least 21 to 23 days of valid daily data for acceptance, depending on the number of days in the month. Hourly ozone, NO_x, NO, and NO_2 data were also corrected for baseline drift. All measurements and data processing procedures were conducted in accordance with documented quality assurance guidelines (Bowers and Taylor 2000).

INSTRUMENTATION

The CHS air monitoring equipment was state-of-the-art at the time the network was established. Ozone was monitored hourly using EPA-approved UV photometric instruments. NO_2 was determined hourly from EPA-approved chemiluminescent instruments measuring NOx and NO; NO_2 was calculated as the difference ($NO_2 = NO_x$ – NO). Automated daily calibration systems were employed with all ozone and NO_x instruments. Hourly PM_{10} mass was initially measured using the TEOM instrument (Patashnick and Rupprecht 1991). The standard operating procedure for the TEOMs included heating the incoming air to 50^oC in order to evaporate water from the aerosol prior to collection. The PM_{10} concentrations measured by TEOMs were subsequently adjusted to account for the possible inadvertent evaporation of ammonium nitrate and organic compounds. Additional instrumentation was installed across the sampling network in 2000 and 2001 to obtain more data about primary emission pollutants. Carbon monoxide analyzers were installed at all of the stations except Alpine, and condensation particle counters (CPC, TSI Model 3022A) were installed at all of the

stations to assess ambient ultrafine particle number concentration (CPC operations were discontinued after 2003). Some of the routine monitoring sites were equipped with $PM_{2.5}$ FRM filter samplers collecting daily, 1-in-3 day, or 1-in-6 day samples starting in 1998. In 2005–2008, continuous $PM_{2.5}$ MetOne BAMs were installed at sites that did not have daily $PM_{2.5}$ FRM sample collections. Also in 2009, the aging PM_{10} TEOMs were replaced with the newer and more reliable MetOne BAMs.

To obtain long-term air monitoring information about various ambient acids across multiple communities for years of study in a cost-effective manner, a novel aerosol/acid sampler (the TWS) was specifically developed for study use in 1991–1992 by Hering and colleagues (1994). This sampler was used to collect time-integrated samples of $PM_{2.5}$ mass, $PM_{2.5}$ chemical components, and selected gaseous acids for consecutive two-week periods throughout the initial decade of CHS. Prior to study deployment, an instrumentation development effort was required, since existing aerosol/acid technologies were designed for 4- to 24-hour sampling intervals, and the use of short-interval sampling was prohibitively expensive for the envisioned decade-long sampling effort. A major effort was undertaken to develop an acceptable, reliable, and useful two-week aerosol/acid sampler. Details of the sampler development and evaluation effort are provided in Lurmann and colleagues (1994) and Hering and colleagues (1994). The TWS was initially designed with three sampling legs (A, B, and C) and subsequently modified in 2000 to incorporate two additional sampling legs (D and E). The standard operating procedures for the new sampler are described in Taylor and colleagues (2003).

Leg A consisted of a single-jet Teflon impactor that removed particles with aerodynamic diameters theoretically larger than 2.5 µm. The Teflon impactor was connected to a carbonate-coated glass honeycomb denuder that collected nitric and hydrochloric acids, and this was followed by a Teflon filter and a carbonate-impregnated quartz filter for fine-particle mass and ions collection and subsequent analysis. Ion chromatography was used to measure nitrate and chloride during subsequent chemical denuder extraction in the laboratory. This allowed for the determination of "ambient strong acids" (nitric acid and hydrochloric acid concentrations). The Teflon filter was analyzed gravimetrically for $PM_{2.5}$ mass and by ion chromatography for $SO_4^=$, NO_3^-, and NH_4^+. The quartz backup filter, coated with sodium carbonate (Na_2CO_3), was used to collect volatilized particulate nitrate, and then analyzed by ion chromatography. The sampler flow rate for this leg of the TWS was 0.4 L/min. $PM_{2.5}$ mass was determined from the sum of the mass measured on the Teflon front filter and the mass of ammonium nitrate determined by nitrate measurements on the backup filter. Total nitrate ion concentration was determined as the sum of front and backup filter nitrate. The total ammonium ion concentration was the sum of the front filter ammonium and the ammonium presumed to be associated with the measured backup filter nitrate (i.e., 18/62 of the backup filter nitrate, based on molecular weight proportionality).

Leg B utilized the filter-pack methodology of Solomon and colleagues (1988) for determination of two gaseous organic acids. Leg B contained sampling cartridge with a Teflon pre-filter to remove particles, followed by two potassium hydroxide–impregnated quartz fiber filters. The material collected on the filters was extracted and analyzed by ion chromatography for formate and acetate to provide ambient "weak acid"

concentrations (formic acid and acetic acid). The acetic acid determination was an upper-limit estimate because this measurement method is known to have positive interference from peroxyacetylnitrate (PAN) (Grosjean and Parmar 1990). CHS Sampler Leg B flow rate was controlled to 0.4 L/min.

The PM sampling train on TWS Leg C contained a baked quartz fiber filter from which organic and elemental carbon (OC and EC, respectively) were determined by thermal/optical transmittance laboratory analyses. The sampler leg flow rate was 1.3 L/min. The sampler leg was comprised of an inlet line (6 mm diameter by 2.5 cm long) and a Teflon filter holder (Savilex). It did not contain a size-selective inlet, yet Salmon and colleagues (2001) demonstrated that Leg C of the sampler collected an effective PM_{10} sample. The laboratory analysis procedure was similar to the NIOSH method (NIOSH 1996) and has been previously described by Birch and Carey (1996).

Leg D was identical to Leg C except that it incorporated a size-selective inlet (Mingchih et al. 1999). Sampling Leg D consisted of an aluminum $PM_{2.5}$ impactor followed by a baked quartz fiber filter from which $PM_{2.5}$ OC and EC were determined by thermal/optical transmittance laboratory analyses. Sample flow rate in this leg of the sampler was 1.3 L/min. The impactor had a conical cavity with a glass fiber substrate that was coated with two drops of mineral oil to collect large particles (Mallinckrodt N.F. white mineral oil #6358 was purposely used so as not to interfere with OC/EC measurements).

Leg E contained a $PM_{2.5}$ impactor followed by a Teflon filter. This sampling leg was used to provide a sample substrate from which elements from aluminum to uranium could be determined using X-ray fluorescence (XRF) analyses. Sampler leg flow rate

was 1.3 L/min. The laboratory analysis method for the XRF analyses undertaken was described by Watson and colleagues (1999).

CONCENTRATION ADJUSTMENTS

PM concentration measured by TEOMs, BAMs, and TWS were adjusted to represent FRM PM concentrations, as best as possible, throughout the decades of study operations. For example, PM_{10} concentrations measured using TEOMs with heated inlets are now known to be biased low, due to the volatilization of ammonium nitrate and certain organic compounds (Allen et al. 1997). To account for the loss of semi-volatile species in the heated TEOM inlets, regression relationships were developed relating the 24-hr average PM_{10} concentrations measured by TEOM to collocated PM_{10} HiVol sampler-based measurements (the FRM for PM_{10} during the first several years of study field operations). Initially, collocated measurements were available at four locations: Rubidoux, Lancaster, Atascadero, and Long Beach. The TEOM adjustment factor was highest (1.25) at Rubidoux, located east of Los Angeles between Mira Loma and Riverside. This adjustment factor was used for all of the low-elevation inland sites in the South Coast Air Basin (SoCAB), including San Dimas, Glendora, Upland, Mira Loma, and Riverside. A TEOM adjustment factor of 1.02 was derived from the Atascadero site data and used to adjust collected data at the cleaner ambient PM_{10} concentration sites, including Santa Maria and Lompoc. A TEOM adjustment factor of 1.11, derived from the Lancaster station data, was used for the other low-to-moderate PM_{10} concentration sites, including Alpine and Lake Arrowhead. TEOM concentrations at Lake Elsinore were adjusted using the Long Beach adjustment

factor (1.18). These factors were periodically updated based on subsequent comparisons to collocated FRM data.

PM concentrations measured with BAMs had the opposite problem to those of TEOMs and tended to read higher than co-located FRM values, mostly due to aerosol water retention. Annual adjustment factors of 0.56 to 0.98 for $PM_{2.5}$ and PM_{10} BAMs were developed from collocated PM FRM data for each monitor and, where feasible, were updated over time.

After the San Dimas air monitoring station was closed in 1996, pollutant concentration estimates for the San Dimas community were estimated from measurements obtained at the Glendora regional monitoring station. The Glendora station was located on the outskirts of Glendora and about 3 km from the nearest freeway. A comparison of one to three years of concurrent monitoring at San Dimas and Glendora indicated that pollutant concentrations were highly correlated, with significant bias (>10%) for NO, NO_2, nighttime ozone, $PM_{2.5}$ EC, and $PM_{2.5}$ nitrate. The average EC and nitrate concentrations were 29% and 19% higher, respectively, at San Dimas than Glendora. Average differences in $PM_{2.5}$ mass and ammonium were less than 10%, so no adjustments were made. The largest bias was in NO concentrations, which were 100% higher on a 24-hour average basis and ~500% higher for the 1-hour morning maximum (on average) at San Dimas compared to Glendora. Much smaller differences were found for NO_2, which was higher at San Dimas than at Glendora, and for nighttime ozone, which was lower at San Dimas than at Glendora. Regressions for each hour of the day were used to estimate San Dimas NO_2 concentrations, based on collected measurement data from Glendora and Azusa. A similar approach was used to estimate

San Dimas ozone concentrations, based on collected ozone measurements at Glendora.

For direct intensive comparison of the TWS, only one CHS community site (north Long Beach, from 1994 to 2001 period) provided a location where daily $PM_{2.5}$ FRM samples were collected next to an operating TWS. Comparison of $PM_{2.5}$ mass data for 19 two-week periods in 2000 and 2001 indicated the TWS $PM_{2.5}$ mass measurements were within ±2.7 μg/m^3 and ±15% of the FRM measurements on average, and were biased by -2.3 μg/m^3 and -12.5% on average. Note that for some periods, the two-week averages of FRM measurements were determined from 13 days rather than 14 days of data, which may have contributed to the bias. Figure A.2 shows that the measurements were well-correlated (r^2 = 0.80), but biased low. This bias to the "low" side was presumably due to loss of volatile compounds over the two-week continuous sampling period. The results of comparisons with FRM data were comparable to those made with lesser-duration intensive sampling study data, such as those from the Southern California Air Quality Study (SCAQS) data set made during the sampler development study in 1993. Also, Motallebi and colleagues (2003) found similar results when they compared two-week average TWS mass concentrations with the mean of ten daily FRM measurements made within two-week periods at Long Beach in 1999 and 2000. Based on these collective analyses, an adjustment factor of 1.17 was applied to all TWS mass data to represent FRM equivalent data.
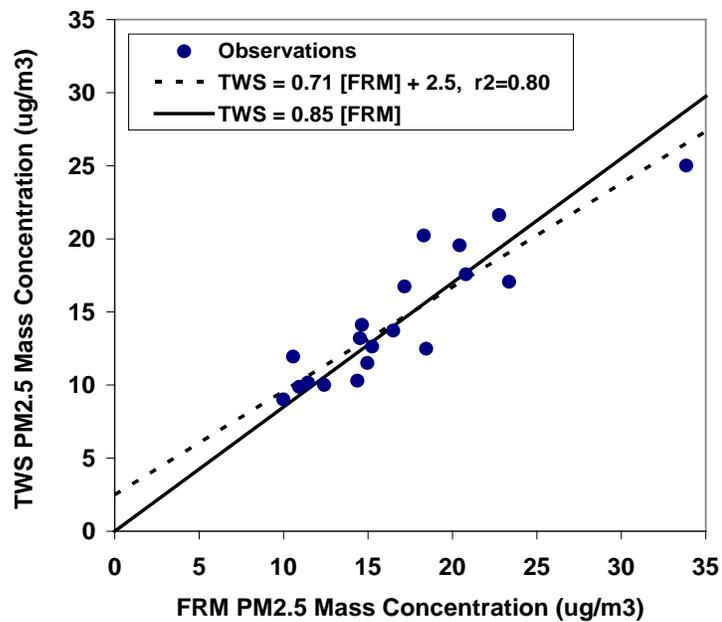
Figure A.2. Comparison of two-week average PM$_{2.5}$ mass from FRM and CHS two-week sampler (TWS) in Long Beach in 2000–2001.

TWO-WEEK SAMPLER PRECISION

In-use precision of the TWS was determined from 13 site-years of collocated data. Two additional samplers were operated at Riverside, Mira Loma, Lake Elsinore, and/or Alpine sampling sites from 1994 to 2001. Figure A.3 illustrates the comparison between the primary and secondary samples collected at all locations. Based on the data that exceeded the detection limits at all sites where data was collected, mean absolute differences in collocated PM$_{2.5}$ mass, nitrate, sulfate, and ammonium concentrations were ±2.0, ±0.7, ±0.2, and ±0.3 µg/m^3 and ±11%, ±10%, ±9%, and ±13%, respectively. The mean absolute differences in collocated acetic acid, formic acid, nitric acid, and hydrochloric acid determinations were ±0.30, ±0.14, ±0.36, and ±0.16 ppb and ±9%, ±12%, ±12%, and ±20%, respectively. For these types of analyses, precision is best determined by the pooled coefficients of variation. The TWS pooled coefficients of variations were 10%, 11%, 11%, and 15% for PM$_{2.5}$ mass, nitrate, sulfate,

and ammonium, respectively, and 8%, 12%, 13%, and 22% for acetic acid, formic acid, nitric acid, and hydrochloric acid, respectively. All the TWS measurements, except those for the hydrochloric acid, met the original design precision goal of ±15%. The ambient concentrations of hydrochloric acid were quite low and variable, and were not used for any health analyses.
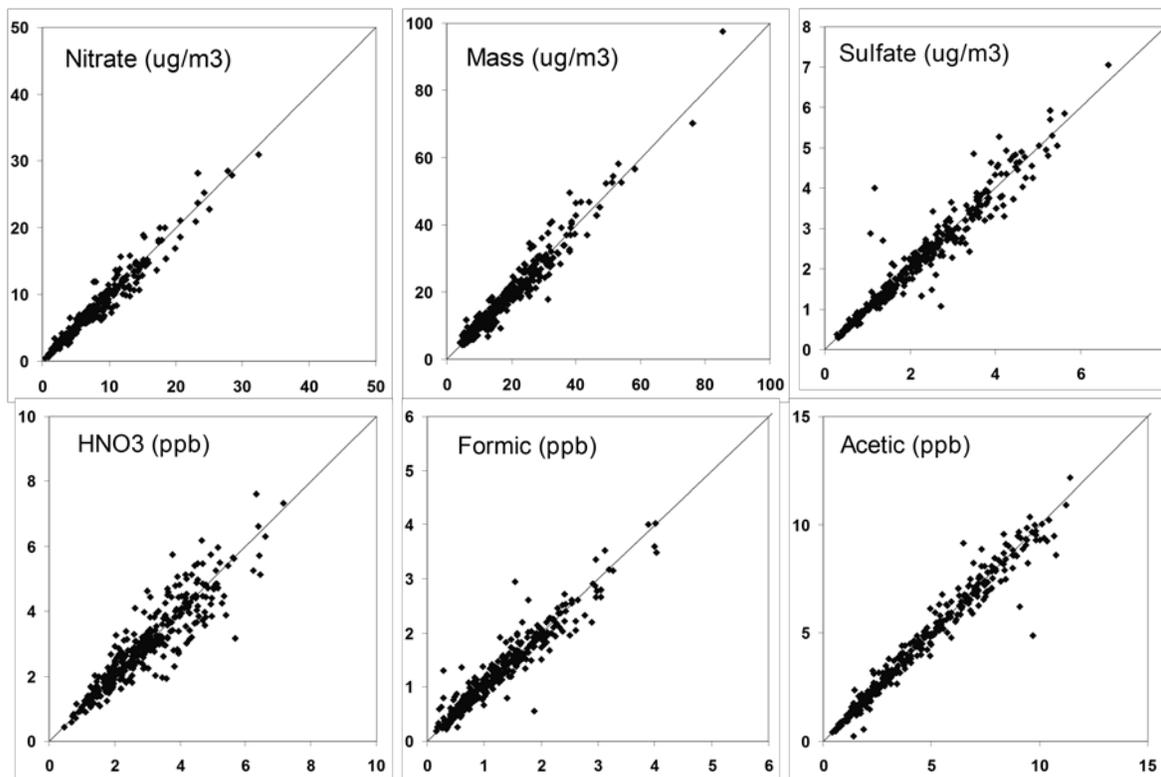

Figure A.3. Comparison of collocated two-week sampler concentrations of PM$_{2.5}$ mass, PM$_{2.5}$ nitrate, PM$_{2.5}$ sulfate, nitric acid, formic acid, and acetic acid from 1995–2001.

PROCEDURES TO FILL IN MISSING DATA

Missing air quality data are an inevitable byproduct of extended field sampling in numerous locations over periods of years. A number of procedures were used to fill in data gaps. First, it is important to note that some of the air data gaps were intentional. For example, the respective monitoring station gas analyzers (for ozone, NO, NO$_2$, and

NO$_x$) typically have one hour of data per day missing (usually in the early morning) because of instrument auto-calibrations. During other sampling times of possible study interest, data gaps may have occurred. When one or two-hour gaps in hourly data occurred, they were filled by linearly interpolating the concentrations from the concentrations at the hour before and the hour after the data gap. This approach was applied to both particle and gas monitoring data.

On much rarer occasions, due to instrument or data logger problems, data gaps ranging from three hours to several months occurred over the 20-year CHS data collection period. In most cases, concurrent data from nearby stations were used during periods of successful data location to develop highly accurate inter-station regressions. When multi-hour or multi-month data gaps did occur, these data regressions from neighboring stations were used to predict concentrations at the CHS stations of interest. However, nearby measurements for the pollutant with missing data were not always available in some communities, especially in the early years (1990s) for NO$_2$, PM$_{2.5}$, and PM$_{10}$. Ozone concentrations were generally available and easily filled using nearby data and regression equations developed from one or more years of data. NO$_2$, PM$_{2.5}$, and PM$_{10}$ were filled from regression of nearby station data when the regressions had r^2>0.50 (most had r^2>0.80).

There were some longer periods without data for certain pollutants at the desired monitor locations. For example, the field study continued for several years without on-site PM$_{2.5}$ data from the Upland and Riverside stations. At those locations, data from nearby monitoring stations were used. Accordingly, Ontario station data was substituted for missing Upland data and Rubidoux and/or Magnolia station data were used in lieu of

available data for Riverside. In a similar manner, long data sampling gaps in Santa

Barbara were filled using regression equation replacement data from the next closest

reporting station in Goleta. Data gaps at Long Beach were filled in using Anaheim data.

In San Dimas, where the station was permanently shut down by school district request,

we found that hourly ozone and NO_2, daily PM_{10}, and two-week $PM_{2.5}$ could be

accurately estimated from routinely collected data at the nearby Glendora and Azusa

monitoring stations (on the basis of one year of concurrent data comparisons, as

described previously). More challenging data gaps occurred for PM and NO_2 at the

higher elevation sites of Alpine and Lake Arrowhead, because the nearest sites with

measurements (often at lower elevations) were not predictive of conditions at the CHS

sites.

The ozone, NO_2, and PM data for the CHS communities had strong seasonal

patterns. In computing annual average concentrations where monthly averages were

missing, estimates of missing monthly averages from the concentrations at the site

during the prior and subsequent year were used to compute annual averages for the

missing time period. This procedure is similar to EPA's approach to using quarterly

averages to obtain valid annual averages because it avoids biasing the results by the

seasonal nature of the missingness.

SUPPLEMENTAL METHODS

STATISTICAL MODELING

General Modeling Approach

For analyzing the effect of annual averaged air pollution concentrations on bronchitic symptoms or lung function, we followed a multi-level modeling strategy described previously (Gauderman et al. 2007; McConnell et al. 2006; Berhane et al. 2004; McConnell et al. 2003; Gauderman et al. 2015). To summarize, a three-level generalized linear mixed model was used in this study, with *t, c, i,* and *j* denoting the community, cohort, subject, and year of visit, respectively. In the first level, we modeled bronchitic symptoms or lung function (denoted by y) as a function of time-dependent covariates $\boldsymbol{Z_{tcij}}$. The $g(.)$ in Equation 1 represents the appropriate link function for the model outcome (logit link for the binary bronchitic symptoms and identity link for the continuous lung function). A_{tci} and B_{tci} represent subject-specific intercepts and slopes, respectively, which were used in the second level models. T_{tcij} represents age, the time variable used for analysis. Spline terms were included in the model to account for non-linear changes in outcome.

$$g(E[Y_{tcij}]) = A_{tci} + B_{tci}T_{tcij} + \boldsymbol{\gamma_1 Z_{tcij}} \qquad [1]$$

In the second-level model (Equations 2a and 2b), the subject-specific intercepts and slopes were regressed against subject-specific covariates, $\boldsymbol{Z_{2a,tci}}$ and $\boldsymbol{Z_{2b,tci}}$. A_{tc} and

B_{tc} represent community and cohort-specific intercepts and slopes, which were used in the third-level models.

$$A_{tci} = A_{tc} + \gamma_{2a}\, Z_{2b,tci} + e_{a,tci} \qquad\qquad [2a]$$

$$B_{tci} = B_{tc} + \gamma_{2b}\, Z_{2b,tci} + e_{b,tci} \qquad\qquad [2b]$$

In the third-level model, community and cohort-specific intercepts A_{tc} from level 2 were regressed as functions of the cohort-specific annual averages of air pollution for each community (Equation 3a). Similarly, community-cohort-specific slopes B_{tc} from level 2 were regressed as functions of the cohort-specific annual averages of air pollution for each community (Equation 3b). Community indicators were also included in the third-level model as adjustments to focus health-effect estimates on secular changes in air pollutions rather than across communities.

$$A_{tc} = \alpha_a + \beta_a \bar{X}_{tc} + \gamma_{3a} Community + e_{a,tc} \qquad\qquad [3a]$$

$$B_{tc} = \alpha_b + \beta_b \bar{X}_{tc} + \gamma_{3b} Community + e_{b,tc} \qquad\qquad [3b]$$

These regression models from the three levels were combined to yield a more efficient mixed-effects model (Equation 4). Note also that the second and third level models on the rate of change (i.e., slope) are only relevant for the lung function outcome.

$g\ (E[y_{tcij}])$

$= (\alpha_a + \beta_a \bar{X}_{tc} + \gamma_{3a} Community + \boldsymbol{\gamma_{2a}}\, \boldsymbol{Z_{2a,tci}} + \boldsymbol{\gamma_1} \boldsymbol{Z_{tcij}})$

$+ (\alpha_b + \beta_b \bar{X}_{tc} + \gamma_{3b} Community + \gamma_{2b}\, Z_{2b,tci}) T_{tcij}$

$+ (e_{b,tc} + e_{b,tci})\, T_{tcij} + e_{a,tc} + e_{a,tci}$

$$[4]$$

Bronchitic Symptoms Modeling

As the main modeling paradigm for analyzing the effect of improvements in air quality on bronchitic symptoms, a three-level logistic linear mixed effects model version of the overall multi-level model specified above (Equation 5) was used in this study (Gauderman et al. 2007; McConnell et al. 2006; Berhane et al. 2004; McConnell et al. 2003). In the first level, we modeled the binary outcome on bronchitic symptoms (denoted by $y = 0\ or\ 1$ as a function of time-dependent covariates $\boldsymbol{Z_{tcij}}$. Here, A_{tci} represent subject-specific adjusted probability of bronchitic symptoms, which were used in the second level models. T_{tcij} represents age (appropriately centered, say at 10 or 15 years of age to focus the inference), the time variable used for analysis. Based on extensive exploratory analysis, a flexible piecewise cubic spline function, denoted by $f\left(T_{tcij}\right) = \delta_0 + \delta_1 T_{tcij} + \sum_k \delta_{1+k} g\left(T_{tcij(k)}\right)$, where $g\left(T_{tcij(k)}\right)$ represents the non-linear components of the spline functions, was included in Equation 5 below with breakpoints at ages of 10 and 15 years to account for non-linear age related trends in bronchitic symptoms (Berhane et al. 2004; DeBoor 1974; Berhane et al. 2000).

$$Logit(P[y_{tcij} = 1]) = A_{tci} + f(T_{tcij}) + \boldsymbol{\gamma_1} \boldsymbol{Z_{tcij}} \qquad [5]$$

In the second-level model (Equation 6), subject-specific intercepts (i.e., adjusted probabilities of bronchitic symptoms) were regressed against subject-specific covariates, $Z_{2,tci}$. A_{tc} represents community and cohort-specific prevalences of bronchitic symptoms, which were used in the third-level models.

$$A_{tci} = A_{tc} + \gamma_2 Z_{2,tci} + e_{tci} \qquad [6]$$

In the third-level, community and cohort-specific adjusted prevalences of bronchitic symptoms A_{tc} from level 2 were regressed as functions of the cohort-specific multi-year means of air pollution for each community (Equation 7). Community indicators were also included in the third-level model as adjustments to focus health-effect estimates on secular changes in air pollution rather than across communities.

$$A_{tc} = \alpha + \beta \bar{X}_{tc} + \gamma_3 Community + e_{tc} \qquad [7]$$

These three regression models were combined to yield a more efficient mixed-effects model (Equation 8).

$$Logit(P[y_{tcij} = 1]) = \alpha + \beta \bar{X}_{tc} + \gamma_3 Community$$
$$+ \gamma_2 Z_{2,tci} + \gamma_1 Z_{tcij} + f(T_{tcij}) + e_{tc} + e_{tci}$$
$$[8]$$

The main parameter of interest in the above logistic mixed-effects model is the within-community across-cohorts effect of improvements of air quality on trends in bronchitic symptoms, β. To conduct a sensitivity analysis on the implicit assumption of homogeneity of the within-community effects across communities, we used a modified version of Equation 7 as follows:

$$A_{tc} = \alpha + b_c \bar{X}_{tc} + \gamma_3 Community + e_{tc}$$
$$b_c = \beta + f_c$$

$$[9]$$

$$\rightarrow A_{tc} = \alpha + \beta \bar{X}_{tc} + \gamma_3 Community + e_{tc} + f_c \bar{X}_{tc} \qquad [10]$$

Here, we assume $f_c \sim N(0, V_f)$ and the test for heterogeneity is based on whether V_f is significantly different from zero. We finally note that our modeling paradigm is consistent with previous models fitted to CHS data (e.g., McConnell et al. 2003) where the interest was on assessing the effects of the within-community yearly fluctuations of pollution levels from the multi-year community means (i.e., $(X_{tcj} - \bar{X}_{tc})$ for a given cohort c), and the between-community differences $\bar{X}_{tc}$ for a given cohort c. In the present analysis, we are essentially decomposing $(X_{tcj} - \bar{X}_c)$ into within-cohort between-years $(X_{tcj} - \bar{X}_{tc})$ and between-cohorts within-community $(\bar{X}_{tc} - \bar{X}_c)$ effects. Our results in the present work are based on β, that is, the effect associated with

$(\bar{X}_{tc} - \bar{X}_c)$. Due to the use of community-specific intercepts, we implicitly account for any between-community effects (including air pollution). In sensitivity analysis, inclusion of $(X_{tcj} - \bar{X}_{tc})$ did not show any significant results, indicating that the effects of the substantial improvements in air quality between cohorts C, D, and E were dominating the effect of yearly fluctuations in air pollution levels.

Lung Function Modeling

We used a mixed-effect linear spline model to investigate the relationship between longitudinal measurements of lung function and change in air quality. The use of a spline model was necessary in order to properly depict the non-linear growth trajectory of lung function during adolescence. Linear splines are piecewise linear functions that are joined smoothly at a pre-specified number of breakpoints, known as knots (DeBoor 1974). Such modeling approaches have been used for lung function trajectories in the Harvard Six Cities Study (Wang et al. 1993) and the Southern California Children's Health Study (Gauderman et al. 2004; Gauderman et al. 2007; Berhane et al. 2004; Berhane et al. 2000; Berhane and Molitor 2008; Gauderman et al. 2015).

To describe the model, let y_{tcij} denote the lung function measurement on subject i at visit j in community t and cohort c, and let T_{tcij} denote the corresponding age at which the measurement was recorded. We define $T_{tcij}{}^* = (T_{tcij} - T_0)/R$ to be age-centered at age T_0 and standardized to range R, which allows us, for example, to focus on mean lung function at age T_0 (e.g., age 15) and growth in lung function over range R (e.g., 4 years from 11 to 15) (Gauderman et al. 2015). Let $\bar{X}_c$ denote the multi-year mean level

of a given pollutant (e.g., NO_2 or $PM_{2.5}$, as shown in Investigators' Report Table 7) for community t in cohort c. Let $\boldsymbol{Z_{tcij}}$ denote a vector of adjustment covariates, which includes both time-dependent (height, height-squared, BMI, BMI-squared, presence of a respiratory illness) and time-independent (community, sex, race, Hispanic ethnicity) variables. In addition, interactions between each of sex, race, Hispanic ethnicity, and community with $T_{tcij}*$ were included to allow for separate growth curves for each sex, race, ethnic origin, and community. The model relating longitudinal lung function measurements to age, pollution, and covariates has the form:

$$y_{tcij} = \alpha_a + \alpha_b T_{tcij}* + \beta_a \bar{X}_{tc} + \beta_b \bar{X}_{tc} \times T_{tcij}* + \delta\phi(T_{tcij}*) + \gamma_a \boldsymbol{Z_{tcij}} + \gamma_b \boldsymbol{Z_{tci}} \times T_{tcij}* + e_{a,tc} + e_{a,tci} +$$

$$(e_{b,tc} + e_{b,tci})T_{tcij}* + e_{tcij} \quad [11]$$

The parameters of primary interest are β_b, which quantifies how air pollution for cohort c in community m affects lung function growth over age range R, and β_a, which quantifies how air pollution affects mean lung function at age T_0. The function $\phi(T_{tcij}*)$ parameterizes the linear spline function (after accounting for inclusion of the $T_{tcij}*$ term in the model) with knots at ages 12, 14, and 16. Specifically, we assume the underlying model relating lung function to age is a four-piece linear spline with knot points at ages 12, 14, and 16 years. This form of spline model was chosen as it does a good job of capturing the growth pattern of both boys and girls during adolescence. This model has five parameters, which can be conceptualized as an intercept and four piece-specific slopes (e.g., the slope before age 12, the slope from age 12 to 14, etc.). However, since our primary interest is in the 4-year growth in lung function from age 11 to 15, it was

more convenient to re-parameterize the model as a single slope (the 4-year growth from

11 to 15) and three additional terms reflecting deviations of the piece-specific slopes

from that 4-year slope. Our parameterization of the spline model allowed us to focus our

inferences on the air pollution effects on lung function level (the intercept, β_a in Equation

11) and overall growth (4-year growth from 11 to 15, β_b in Equation 11), while

accounting for the underlying nonlinearity in adolescent lung growth in both boys and

girls. A fixed effect for community is included as part of the adjustment covariates to

focus health-effect estimates and inferences on temporal changes in air pollution (rather

than cross-community comparisons). Random effects for both level and growth are

included at the community * cohort level ($e_{a,tc}$,$e_{b,tc}$) and individual level ($e_{a,tci}$,$e_{b,tci}$), with

an overall residual (e_{tcij}) that captures deviations of each observed lung function

measurement from the fitted model. The HPMIXED procedure in SAS (Version 9.4, SAS

Institute Inc., Cary, NC) was used to fit the models.

REFERENCES

Allen G, Sioutas C, Koutrakis P, Reiss R, Lurmann FW, Roberts PT. 1997. Evaluation of the TEOM method for measurement of ambient particulate mass in urban areas. J Air Waste Manag Assoc 47:682–689.

Berhane K, Gauderman WJ, Stram DO, Thomas DC. 2004. Statistical issues in studies of the long term effects of air pollution: The Southern California Children's Health Study. Stat Sci 19:414–449.

Berhane K, McConnell R, Gilliland F, Islam T, Gauderman WJ, Avol E, et al. 2000. Sex-specific effects of asthma on pulmonary function in children. Am J Respir Crit Care Med 162:1723–1730.

Berhane K, Molitor NT. 2008. A Bayesian approach to functional-based multilevel modeling of longitudinal data: Applications to environmental epidemiology. Biostatistics 9:686–699.

Birch ME, Carey RA. 1996. Elemental carbon-based method for monitoring occupational exposures to diesel exhaust. Aerosol Science and Technology 25:221–241.

Bowers K, Taylor C. 2000. Quality Assurance Program Plan. Epidemiologic investigation to identify chronic health effects of ambient air pollutants in Southern California (Phase III-longitudinal study). Part iia. Air quality measurements. Report prepared by the California Air Resources Board, September.

DeBoor C. 1974. A Practical Guide to Splines. Springer-Verlag:New York.

Gauderman W J, Avol E, Gilliland F, Vora H, Thomas D, Berhane K, et al. 2004. The effect of air pollution on lung development from 10 to 18 years of age. N Engl J Med 351:1057–1067.

Gauderman W J, Urman R, Avol E, Berhane K, McConnell R, Rappaport E, et al. 2015. Association of improved air quality with lung development in children. N Engl J Med 372:905–913.

Gauderman W J, Vora H, McConnell R, Berhane K, Gilliland F, Thomas D, et al. 2007. Effect of exposure to traffic on lung development from 10 to 18 years of age: A cohort study. Lancet 369:571–577.

Grosjean D, Parmar SS. 1990. Interferences from aldehydes and peroxyacetyl nitrate when sampling urban air organic acids on alkaline traps. Environ Sci Technol 24:1021–1026.

Hering SV, Taylor CA Jr, Anderson K. 1994. Design and implementation of a two-week sampler for fine particle ions and gaseous acids. In: Fourth International Aerosol Conference, Los Angeles, CA.

Lurmann FW, Roberts PT, Main HH, Hering SV, Avol EL, Colome SD. 1994. Phase II report Appendix A: Exposure assessment methodology. Final report prepared for the California Air Resources Board, Sacramento, CA, by Sonoma Technology, Inc., Santa Rosa, CA, Aerosol Dynamics, Inc., Berkeley, CA, USC School of Medicine, Los Angeles, CA, and Integrated Environmental Services, Irvine, CA, STI-92340-1423-FR, October.

McConnell R, Berhane K, Gilliland F, Molitor J, Thomas D, Lurmann F, et al. 2003. Prospective study of air pollution and bronchitic symptoms in children with asthma. Am J Respir Crit Care Med 168:790–797.

McConnell R, Berhane K, Yao L, Jerrett M, Lurmann F, Gilliland F, et al. 2006. Traffic, susceptibility, and childhood asthma. Environ Health Perspect 114:766–772.

Mingchih C, Kim S, Sioutas C. 1999. Experimental studies on particle impaction and bounce: Effects of substrate design and material. Atmospheric Environment 33:2313–2322.

Motallebi N, Taylor CA Jr,Turkiewicz K, Croes BE. 2003. Particulate matter in California: Part 1— Intercomparison of several $PM_{2.5}$, $PM_{10-2.5}$, and PM_{10} monitoring networks. J Air Waste Manag Assoc 53(12): 1509–1516.

NIOSH. 1996. Elemental carbon (diesel exhaust). In: NIOSH Manual of Analytical Methods. National Institute of Occupational Safety and Health:Cincinnati, OH.

Patashnick J, Rupprecht EG. 1991. Continuous PM-10 measurements using the tapered element oscillating microbalance. J Air Waste Manag Assoc 41:1079–1083.

Salmon LG, Mertz KA, Mayo PR, Cass GR. 2001. Determination of elemental carbon, organic carbon, and source contributions to atmospheric particles during the Southern California Children's Health Study: Part A; 98–320; California Air Resources Board: Sacramento, CA.

Solomon PA, Fall T, Larson SM, Lin P, Vasquez F, Cass GR. 1988. Acquisition of acid and aerosol concentration data for use in dry deposition studies in the South Coast Air Basin. Vol. I. Report prepared for the California Air Resources Board, Sacramento, CA by California Institute of Technology, Environmental Quality Lab Report No. 25, Contract No. A4-144-32.

Taylor CA. 2003. S.O.P. 001, standard operating procedure for the two-week sampler. Prepared for the California Air Resources Board, Research Division, Sacramento, CA. Wang X , Dockery DW, Wypij D, Gold DR, Speizer FE, Ware JH, et al. 1993. Pulmonary function growth velocity in children 6 to 18 years of age. Am Rev Respir Dis 148:1502–1508.

Watson JG, Chow JC, Frazier CA. 1999. X-ray fluorescence analysis of ambient air samples. In: M. Creatchman (ed). Elemental Analysis of Airborne Particles. Gordon and Breach Publishers:Newark,NJ.