United States
Environmental Protection
Agency

Health Effects Research
Laboratory
Research Triangle Park NC 27711

EPA-600 1-79-034
August 1979

Research and Development

EPA

600179034

Model for Measuring the Health Impact from Changing Levels of Ambient Air Pollution

Mortality Study

EP 600/1
79-034

RESEARCH REPORTING SERIES

Research reports of the Office of Research and Development, U.S. Environmental Protection Agency, have been grouped into nine series. These nine broad categories were established to facilitate further development and application of environmental technology. Elimination of traditional grouping was consciously planned to foster technology transfer and a maximum interface in related fields. The nine series are:

1. Environmental Health Effects Research
2. Environmental Protection Technology
3. Ecological Research
4. Environmental Monitoring
5. Socioeconomic Environmental Studies
6. Scientific and Technical Assessment Reports (STAR)
7. Interagency Energy-Environment Research and Development
8. "Special" Reports
9. Miscellaneous Reports

This report has been assigned to the ENVIRONMENTAL HEALTH EFFECTS RESEARCH series. This series describes projects and studies relating to the tolerances of man for unhealthful substances or conditions. This work is generally assessed from a medical viewpoint, including physiological or psychological studies. In addition to toxicology and other medical specialities, study areas include biomedical instrumentation and health research techniques utilizing animals — but always with intended application to human health measures.

This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161.

EPA-600/1-79-034
August 1979

MODEL FOR MEASURING THE HEALTH IMPACT FROM CHANGING LEVELS OF AMBIENT AIR POLLUTION: MORTALITY STUDY

BY

Tsukasa Namekata, Bertram W. Carnow,
Domenic J. Reda, Eileen B. O'Farrell and James R. Marselle
Occupational & Environmental Medicine Program
School of Public Health

University of Illinois at the Medical Center
Chicago, Illinois 60680

68-02-2492

Dr. Wilson Riggan

Health Effects Research Laboratory
U.S. Environmental Protection Agency
Research Triangle Park, North Carolina 27711

HEALTH EFFECTS RESEARCH LABORATORY
OFFICE OF RESEARCH AND DEVELOPMENT
U.S. ENVIRONMENTAL PROTECTION AGENCY
RESEARCH TRIANGLE PARK, NORTH CAROLINA 27711

DISCLAIMER

This report has been reviewed by the Health Effects Research Laboratory, U.S. Environmental Protection Agency, and approved for publication. Approval does not signify that the contents necessarily reflect the views and policies of the U.S. Environmental protection Agency, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

late and specific types of mortality, a percentage decrease in the age-adjusted death rates was estimated when a 25 percent reduction in TSP, which is almost equivalent to the percentage reduction in TSP in Chicago for the period 1970-75, was applied to the models developed. The age-adjusted death rate for non-accidental causes would be decreased by 5.36% (54.65 deaths per 100,000 persons) in Chicago. A percentage decrease in the death rates by cause was estimated to be 8.82% (all heart diseases), 6.42% (ischemic heart disease), 16.95% (other heart disease), 9.39% (diabetes mellitus), 20.13% (cirrhosis of the liver), 26.16% (emphysema) and 6.47% (other non-accidental causes).

In multiple regression analysis for the daily time-series study, the dependent variables were the number of daily non-accidental deaths and the number of daily deaths due to heart disease throughout the city of Chicago. The independent variables were (1) pollutants (TSP, SO_2 and TSPxSO_2), (2) climatological variables (daily average temperature, wind speed, precipation, snow fall, humidity, sunshine and sky cover), and (3) day-of-week variables as dummy variables. Models developed in daily analysis imply that there would be possible acute effects of daily air pollution concentrations (both SO_2 and TSP, in addition to their interaction) on daily mortality changes (both all non-accidental causes and heart diseases), controlling for weather and day-of-week effects. Models for daily non-accidental deaths could be affected by levels of SO_2, TSP and their interaction on the day of death, and levels of SO_2 and an interaction between SO_2 and TSP on the third and the sixth day prior to death occurence. Models for heart disease indicate that the number of daily deaths caused by heart disease could be affected by levels of SO_2, TSP and their interaction on the day of death onset, levels of SO_2 and an interaction between SO_2 and TSP on the third day prior to death onset, and levels of SO_2 on the sixth day prior to death onset.

This work was submitted in fulfillment of Contract No. 68-02-2492 by the University of Illinois, the School of Public Health, under the sponsorship of the U.S. Environmental Protection Agency. This report covers a period from February 15, 1976 to October 1, 1978, and which was completed as of March 8,1979.

CONTENTS

CONTENTS (continued)

FIGURES

TABLES

ACKNOWLEDGEMENT

The authors would like to give grateful appreciation to Mr. Stephen Goranson, Chief, Air Data Analysis Section and Mr. Laurence Lehrman, Statistician, both of the U.S. EPA, Region V, for their co-operation and assistance in providing data for study analysis.

Mr. James Herman, Superintendent of the Technical Services Division, at the City of Chicago Department of Environmental Control and his associates were generous in providing aerometric data monitored under the auspices of the city.

The authors would like to express our appreciation to Ms. Sharon Kawasaki, Ms. Kiyoka Koizumi, Ms. Elaine Breck and Mr. Norman Iverson for their assistance in data processing.

Also, the services of the Illinois Department of Public Health are greatly appreciated for their providing death certificate tapes and other related information.

Sincere appreciation is expressed to Ms. Carol Albor, Ms. Alberta Braden and Ms. Karen Komar for their secretarial assistance.

A very special acknowledgement is given to Dr. Wilson Riggan, EPA Project Officer, for his untiring efforts, cooperation, and patience with this project.

SECTION 1

INTRODUCTION

The United States Congress passed the Clean Air Act in 1967 and its amendments in 1970 to protect and enhance the quality of the nation's air resources while promoting the public health and welfare and the productive capacity of its population, reducing harmful emissions, and ensuring that air pollution problems will, in the future, be controlled in a systematic way. Also, in 1970 the city of Chicago passed an ordinance which virtually banned coal and garbage burning in individual households and businesses.

Such legislative efforts to control air pollution led to significant decreases in the amount of some pollutants in the city of Chicago; suspended particulate levels dropped 26 percent, and sulfur dioxide levels were cut by 50 percent between 1970 and 1975. However, carbon monoxide and ozone levels have remained high, in spite of Chicago's voluntary auto emission control program.

Coincidentally, the city of Chicago has also experienced a 13 percent decrease in the age-adjusted death rate from all causes, particularly an 18 percent decrease from heart disease, for the years 1970-75. These events led to asking the questions, "Is a recent mortality decline in the city of Chicago caused by a decrease in the amount of major air pollutants such as suspended particulate and sulfur dioxide?" or "How much of the reduction in mortality is the result of lower concentration of the individual pollutants in the city of Chicago?"

To answer these questions, linear models were developed to quantitate the partial contribution of major air pollutants (total suspended particulate and sulfur dioxide) to mortality, controlling for other related factors. The present study consists of two parts; (1) cross-sectional analysis to examine chronic effects of air pollution on mortality, and (2) daily time-series analysis to examine acute effects of air pollution on mortality.

SECTION 2
CONCLUSIONS

The results from cross-sectional analysis indicated that there would be possible chronic effects of air pollution, especially total suspended particulate (TSP), on mortality in the city of Chicago. Controlling for the factors of environmental conditions, income levels and education levels throughout 76 community areas, significant associations were found between total suspended particulate and the age-adjusted death rates for total non-accidental causes, heart disease (both ischemic and non-ischemic or other), diabetes mellitus, cirrhosis of the liver, emphysema and other non-accidental causes, and between TSPxSO_2 (an interaction between the two pollutants) and the age-adjusted death rate for emphysema. No significant associations were observed between any of the pollutants included in the study and the age-adjusted death rates for malignant neoplasms (digestive organs and peritoneum, respiratory systems, and genito-urinary organs), cerebrovascular disease, arteriosclerosis, other circulatory disease, pneumonia and influenza, and congenital anomalies and diseases of early infancy. No significant relationship was found between existing SO_2 levels and any cause of death, although a regression coefficient of SO_2 was close to the significant level at $p<.10$ in the model for emphysema.

Based on the significant associations between total suspended particulate and specific types of mortality, a percentage decrease in the age-adjusted death rate was estimated when a 25 percent reduction in TSP, which is almost equivalent to the percentage reduction in TSP in Chicago for the period 1970-75, was applied to the models developed. The age-adjusted death rate for all non-accidental causes would be decreased by 5.36% (54.65 deaths per 100,000 persons) in Chicago. A decrease in the death rate by cause was estimated to be 8.82% (all heart diseases), 6.42% (ischemic heart disease), 16.95% (other heart disease), 9.39% (diabetes mellitus) 20.13% (cirrhosis of the liver), 26.16% (emphysema) and 6.47% (other non-accidental causes).

Models developed in daily analysis also imply that there would be possible acute effects of daily air pollution concentrations (both SO_2 and TSP, in addition to the interaction) on daily mortality changes (both all non-accidental causes and heart disease), controlling for weather changes and day-of-week effects. Models for daily non-accidental deaths indicate that the number of daily non-accidental deaths could be affected by levels of SO_2, TSP and their interaction on the day of death, and levels of SO_2 and an interaction between SO_2 and TSP on the third and the sixth days prior to death occurence. Based on the model for the day of death onset, it is estimated that a 25 percent reduction in daily levels of each pollutant would decrease daily non-accidental deaths by 1.815% (due to SO_2), 2.045% (due to TSP) and 0.867% (due to an interaction between SO_2 and TSP) in the city of Chicago.

Models for heart disease indicate that the number of daily deaths caused by heart disease could be affected by levels of SO_2, TSP and their interaction on the day of death onset, levels of SO_2 and an interaction between SO_2 and TSP on the third day prior to death onset, and levels of SO_2 on the sixth day prior to death onset. Based on the model for the day of death onset, it is estimated that a 25 percent reduction in daily levels of each pollutant would decrease daily deaths from heart disease by 1.717% (due to SO_2), 2.048% (due to TSP) and 0.940% (due to an interaction between SO_2 and TSP) in the city of Chicago.

SECTION 3

RECOMMENDATION

To further examine chronic effects of air pollution on mortality, it is recommended that:

1. Regression models be developed according to age-groups, to identify a high risk age-group.
2. Regression models be developed according to sex to examine if air pollution affects male's mortality differently from female's.
3. The community population used in the study be re-examined whenever reliable population estimates by age are available at the community level in the mid-year between 1970 and 1980.
4. Community exposure levels of air pollution estimated by the simple interpolation method be evaluated if a better estimation method is developed in the future.
5. To control for occupational exposure, occupation code be obtained from the original death certificates if possible and feasible.

In regard to the daily analysis, it is recommended that regression models be developed according to seasons, because disease and pollution concentration patterns might be different from one season to the next.

The models developed in both cross-sectional and daily analyses can be validated in two ways: (1) replicating the study by using data covering different years (e.g., 1976-78), and (2) replicating the study by using data from a different geographical location (e.g., Gary-Hammond area in Indiana which is heavily industrialized).

Finally, it is recommended that continued effort be made to reduce levels of particulate concentrations, especially respirable particulate, in order to further reduce its impact on health.

SECTION 4

BACKGROUND INFORMATION

There is little doubt among scientists that there is in fact a relationship between health and environmental contamination from air pollution. However, the depth and degree of this association is strongly in question. The scientific community hesitates to wholly accept proof of this relationship, due, in part, to lack of proper controls for important factors affecting health (among them, age, smoking, income level, race and climate). Additionally, the sample size utilized is so small as not to lend confidence to tabulated results. Often measurements of health indices maintain 'built-in' inadequacies, whether it be an index of morbidity or mortality.

This document addresses itself to the relationship between death and air pollution. A review follows below of literature pertaining to cardiovascular and respiratory mortality, both domestic and international in its scope.

CROSS-SECTIONAL STUDIES

Respiratory Disease Mortality

Utilizing data from upstate New York, Winkelstein and co-authors (1967)[1] compared 21 areas surrounding Buffalo, New York for levels of air pollution (suspended particulate), income level and the mortality rate for chronic respiratory disease (inclusive of asthma, emphysema, bronchitis, pneumonia and bronchiectasis). The statistical methodology employed was cross-tabulation. Various factors were controlled for throughout the study: some physical characteristics, median family income, number of years of school completed, percentage of laborers in the work force; the death rates were age-sex-race specific. No personal factors, however, were included. Based on each economic level, (areas 1 through 5), a trend was established between pollution and mortality, whereas mortality increased by 100 percent in white males (50-59 yrs.) ranging from pollution level 1 to pollution

level 4.

Using similar methods, a series of studies were performed by Zeidberg, Horton, and Landau (1967)[2] throughout areas of Nashville, Tennessee. The authors contrasted measurements of air pollution and socio-economic status to health. A quartet of pollution indices were utilized: sulfation (sulfur trioxide), soiling (concentration of haze and smoke), dustfall and sulfur dioxide conc. (24 hr.). Utilizing cross-tabulation with mortality from bronchitis and emphysema with these air pollutant levels and income class, no association was found. However, the authors found that total respiratory disease mortality, as before sex-age-race adjusted, was directly related to the degree of sulfation and soiling. Neither sulfur dioxide level - (24 hr.) or dustfall were significantly related to total respiratory disease mortality. Mortality rate differences were lower in women than in men, as were differences in whites versus non-whites. Those socioeconomic variables utilized (occupational level, schooling, median family income and domestic overcrowding) were unable to explain the recorded associations.

Lepper and co-authors (1969)[3] in Chicago reported that, when controlling for socioeconomic class (median income, education and unemployment), total respiratory deaths varied with the concentration of sulfur dioxide across the city, using cross-tabulations. No other health factors were included as variables.

Between 1929-1930, Mills (1943)[4] conducted a study comparing wards in Pittsburgh and in Cincinnati, reporting significant correlations between pollution (sootfall) and pneumonia mortality for white males. He analyzed the correlation to be 0.47 in Pittsburgh and 0.79 in Cincinnati. He also stated that areas of higher altitude had a lower death rate. In this study, no socio-economic variables were considered or controlled; also excluded were potential environmental or personal differences between the two cities. In a subsequent study in Chicago, Mills (1952)[5] investigated sex-age-race specific pneumonia death rates. He discovered that those death rates were always higher in the more polluted areas of the city, using sootfall and sulfur dioxide as a measurement. The youngest age group studied, that is, thirty to thirty-nine year olds, displayed the greatest differences. As in his earlier study, no other socioeconomic

variables were controlled.

Various mortality studies have been conducted over the past decade correlating respiratory mortality with air pollution in the United Kingdom. Stocks (1959)[6] found a significant correlation between sex-specific death rates from bronchitis and a deposit index and smoke, after controlling for population density. Another study by Stocks (1960)[7] found a significant correlation between smoke density and bronchitis death rates, when controlling for both population density and social class. Ashley (1967)[8] also found a positive correlation while controlling for population density between bronchitis mortality (combining both sexes) and smoke and sulfur dioxide. In a later study, which controlled for social class, population density and type of town, Ashley (1969)[9] found a significant positive association between air pollution (smoke) and the male bronchitis death rate. No personal factors were given consideration.

Gardner, Crawford and Morris (1969)[10] continued the investigation of air pollution and bronchitis mortality. They examined death rates in sixty-one boroughs in England and Wales, using five independent variables in a linear multivariate regression analysis. The five variables were the following: social score (the first principle component of nine social indices); air pollution (coal bought for domestic consumption); latitude; water calcium and rainfall. They incorporated data for four age-sex rates (ages forty-five to sixty-four and sixty-five to seventy-five) for two time periods. For bronchitis mortality they found statistically significant effects of air pollution for each of the four age-sex categories in six of the eight data sets.

Buck and Brown (1964)[11] studying certain areas of England and Wales discovered that smoke and sulfur dioxide concentrations were significantly and positively associated with sex-specific bronchitis mortality. Persons per acre, as well as a social index defined as the proportion of unskilled workers among adult males were used as socioeconomic controls. Although current smoking habits were included in the model, no effects were exhibited.

Wicken and Buck (1964)[12] studied Northeastern England with attention given to six specific areas based on urban versus rural composition. Bronchitis mortality rates were higher in these two urban areas as compared with the four rural areas, facts not fully accounted for by age composition,

smoking habits or social class (which was based on occupation). The urban district did have both the highest bronchitis death rate and air pollution level (smoke). The authors do state that the mortality from bronchitis is more closely associated with air pollution than with personal smoking habits.

Later in 1967, Buck and Wicken[13] utilized a multiplicative model to explain bronchitis mortality rates in Northern Ireland among males. Considering urban-rural residence, in addition to the combined effects of family history of bronchitis plus smoking, they discovered that their model fit the data closely and that increased mortality was significantly associated with increased urbanization. Two similar models were used to define the male death rates from bronchitis in Northeastern England. One model included factors for smoking habits and urban-rural residence; the second substituted social class, as defined by occupation, for smoking habits. Although the data was analyzed separately, urbanization was related positively to bronchitis mortality when controlling for other factors.

Daly (1959)[14] reported simple correlations of 0.60 for pneumonia mortality and consumption of domestic fuels and 0.52 for pneumonia and industrial fuels. Tuberculosis mortality was included, yielding correlations of 0.59 and 0.22 respectively. Although simple correlation of both types of mortality using four socio-economic indices were presented (social class, overcrowding, population density and education) no other methodologies (multivariate analysis, e.g.) were utilized.

Collins, Kasap and Holland (1971)[15] working in England and Wales investigated childhood mortality from respiratory causes. Studying the time frame of 1950-1953 and 1959-1963, they discovered that infants less than one year old living under crowded circumstances were the highest risk population from bronchitis and tuberculosis. They also studied mortality from all causes in the period 1958-64, with regard to industrial pollution, sulfur pollution (from stations), domestic pollution, population density, social class, overcrowding and education level. For these children under age one, all variables (save sulfur pollution) were significantly correlated with total mortality and mortality from bronchopneumonia and total respiratory diseases. As children grew older, the association grew weaker.

Japan has also added its research to the or-going bank of mortality

data. Toyama (1964)[16], looking at twenty-one districts in Toyko, discovered a significant correlation between bronchitis mortality and monthly dustfall. However, the study was limited by the author by not accounting for variables other than pollution which might have varied across the districts.

Cardiovascular Disease Mortality

As a function of the Nashville Air Pollution Study, Zeidberg and coauthors (1976)[17] looked at cardiovascular mortality and found an association between age-adjusted death rates for middle class and particulate pollution. When sex was analyzed, females, but not males, presented a consistent association. Rates were higher for each pollution level among non-whites than among whites within the same socioeconomic group.

Enterline and coauthors (1960)[18] had chosen to study the role of urbanization in heart disease. They discovered a higher heart disease mortality rate among forty-five to sixty-four-year-old whites in national center-city counties than in non-metropolitan counties. In metropolitan areas with central cities, males registered a heart disease death rate 37 percent higher than their counter parts in non-metropolitan counties; white females registered a CHD death rate 46 percent higher than those females living in more rural areas.

Other authors have recently performed similar studies analyzing an urban component of health status. Sauer and coauthors (1966)[19] detailed cardiovascular mortality in the southern states of North Carolina and Georgia. They found that white males (aged forty-five to sixty-four and sixty-five to seventy-four) exhibited higher age-adjusted mortality in metropolitan areas than in non-metropolitan areas. Again, the study was limited by the lack of consideration given to other factors. In another similar study, Friedman (1967)[20] correlated mortality rates from coronary heart disease in white males aged forty-five to sixty-four with a proportionate amount of those living in urban areas. For thirty-three states, the simple correlation was 0.79. The partial correlation was 0.67 when cigarette consumption was held constant.

In the United Kingdom, specifically England and Wales, Gardner, Crawford and Morris (1969)[21] could not maintain any constancy in the mortality rate from cardiovascular disease regression analyses. A positive, significant association did exist between males aged forty-five to sixty-four

and air pollution, using domestic coal consumption as an index. Using two different frames, the authors explained air pollution to be the most powerful variable, yielding values of R^2=0.80 and R^2=0.84, respectively. However, rates for males aged sixty-five to seventy-four detailed a statistically non-significant (negative) relationship. Results for females were also statistically non-significant.

EPISODE STUDIES

Greenburg and coauthors (1962)[22] investigated a period of increased sulfur dioxide and smokeshade levels, due to a high stagnant air mass in New York City in November, 1953. The values for these air pollutants were decidedly higher than the average range. Using analysis of variance, they compared this period against six control years, 1950-1956, while assuming a three day lag for the effects of the pollutants. They did find a statistically significant increase (at .05 level) for the number of deaths.

Again Greenburg and associates (1976)[23] investigated an air pollution episode in New York City from January 29 to February 12, 1963, comparing the number of deaths then with a similar time frame in 1961, 62, 64 and 65. These control years were marked by the presence of influenza and cold weather, but not by sulfur dioxide or smokeshade pollution. For the two week period under consideration, they discovered an excess of 200 to 400 deaths which they subsequently attributed to air pollution. The specific causation of these deaths for persons aged over forty-five was attributed to influenza, vascular lesions, cardiac disease and "all others". There were no relevant increases in deaths due to accidents, suicide and homicides.

Gore and Shaddick (1958)[24] examined sex-specific mortality (both total and five categories) during episodes of fog and high air pollution from sulfur dioxide and smoke in London, England. They discovered that during these episodes, critical levels of pollution reaching four times the winter average correlated with excess mortality. For a two year period encompassing these episodes, no significant association between mortality and sulfur dioxide/smoke exposure was displayed. When length of residence in London was added to these pollution indices, significant correlations were found between both sulfur dioxide and smoke and female and male bronchitis mortality.

Glasser, Greenburg and Field (1967)[25] examined mortality (as well as morbidity) from an air pollution episode in New York City which occurred November 23 to November 25, 1976 in concurrence with the Thanksgiving Holiday weekend. With the occurrence of higher than normal air pollution levels, daily deaths from all causes rose to higher than expected levels, remaining high for the following week. There was charted an excess of 24 deaths more per day than during a control period. Thus,over the week there was a total of 168 excess deaths.

Though not described in terms of episodic occurrences of air pollution, two recent studies detailing daily mortality merit investigation. The study by Glasser and Greenburg (1971)[26] investigated daily deaths for the time frame 1960 through 1964 (with the omission of April to September) in New York City. They analyzed deviations from the daily number of deaths against a five year control, by utilizing air pollution variables (smoke-shade and sulfur-dioxide) and weather measurements (wind speed, sky cover, rainfall and temperature deviation from the normal). Using a variety of statistical procedures (cross-tabulation and regression analysis), they found a relationship between daily mortality and sulfur dioxide air pollution. The authors also investigated mortality in terms of the variation from a fifteen day moving average. This measurement did provide cycle in the data and lag effects from pollution, but the authors did not forward an explanation. In this portion of the study, no lagged variables were employed.

Later in 1972, Schimmel and Greenburg (1972)[27] created an additional time-series study of New York City. They included a wide data base: they observed daily total mortality, in addition to nine disease-specific mortality rates from 1963-1968, two air contaminant variables (24-hr. sulfur dioxide and smokeshade readings), and several weather variables (precipitation, wind speed, max./min. humidity and max./min. temperature). Their major analysis was to regress daily mortality on same day pollution levels and air pollution levels on previous days. The authors declared that should air pollution in New York City be reduced to zero, there would be from 18 to 36 fewer deaths each day on the average (the range based on the different pollution variables which were tried). Looking at individual

effects from each of the two pollutants, they estimated that 80 per cent of the excess deaths could be attributed to smokeshades, and only 20 per cent to sulfur dioxide.

SECTION 5

THE METHOD

CROSS-SECTIONAL ANALYSIS

Age-Adjusted Death Rates

The Illinois deaths tapes provided by the Illinois Department of Public Health were used to obtain mortality information for the period 1971-75. Specifically, this information is: the date of death, a sex/race code, the county and subdivision (or community area) of residence of the deceased (which are referenced by tables also provided by the aforementioned agency), the age of the deceased and the underlying cause of death which is a 3-digit code referenced by the Eighth Revision International Classification of Diseases.

The first step was to create a Chicago-only mortality file, discarding all other records. Next, this file was resorted by community area (CA) and date to be used to calculate daily mortality over the years 1971-75. Then 38 different causes of death were selected and grouped into 12 major categories. Some of these categories were later subdivided so that a final total of 17 causes of death were studied. See Table 5.1 for a list of these causes and their corresponding ICD codes. The daily mortality totals were cause-specific with respect to 17 different causes of death, and age specific with respect to 11 age groups (0-4, 5-9, 10-14, 15-19, 20-24, 25-34, 35-44, 45-54, 55-64, 65-74, and 75+). For each day there were 17 x 11 = 187 totals; these were stored by CA and date/ These daily totals were used to produce 5-year totals by CA, retaining the above mentioned specificity. These, then, have 187 x 76 totals for the 5 year period, where 76 is the number of community areas in Chicago.

The totals were then used to determine CA 5-year average death rates, age-specific for each of the 17 disease categories. 1970 population figures for the corresponding 11 age groups in each CA were obtained by combining

Table 5.1
Selected Causes of Deaths for Cross-Sectional Analysis

Causes	ICD code # (8th revision)
(1) All causes excluding accidents, homicides & suicides	
(2) Malignant neoplasms	140-163, 170-174, 180-209
a. Neoplasms of digestive organs and peritoneum	150-159
b. Neoplasms of respiratory systems	160-163
c. Neoplasms of genito-urinary organs	180-189
(3) Heart disease	393-398, 402, 404, 410-414, 420-429
a. Ischemic heart disease	410-414
b. other heart diseases	393-398, 402, 404, 420-429
(4) Cerebrovascular disease	430-438
(5) Arteriosclerosis	440
(6) Other circulatory diseases	390-392, 441-458
(7) Diabetes mellitus	250
(8) Cirrhosis of liver	571
(9) Pneumonia and influenza	470-474, 480-486
(10) Emphysema	492
(11) Congenital anomalies and diseases of early infancy	740-779
(12) All others not included in (2) - (11) and excluded accidents, homicides and suicides	

those from all census tracts in that CA.[28] These population figures were used as yearly populations during the study period (1971-75). The denominators for death rate calculations were five times the age-specific populations in each community area. Since no reliable yearly population estimates are available at the CA level during the study period, this study assumes little population change in the 5 year period. Age and cause specific death rates were calculated in the usual manner (death totals divided by population estimates).

The last step was to calculate 5-year age-adjusted death rates from the 5-year average of age-specific rates for each CA. The following procedure was used:

$$S(i,\ t) = \frac{D(i,t)}{R(i,t)}$$

$$A(t) = \frac{\sum_{i=1}^{n} S(i,t) * P(i)}{\sum_{i=1}^{n} P(i)} \cdot 100{,}000$$

where:

S(i,t) = A proportion of deaths to the number of persons in age group i over t years

A(t) = The age-adjusted death rate over t years

D(i,t) = The number dead in age group i over t years

R(i,t) = The number in age group i in the population at risk over t years

P(i) = The number in age group i in the standard population in the base year (the 1970 total U.S. population multiplied by 5 in this study)

t = The number of years observed (t=5 in this study)

i = A particular age group

n = The number of age groups (n=11 in this study)

Scores Measuring Environmental and Socioeconomic Conditions in Community Areas

In order to quantitate the relationship between air pollution and mortality, it is essential to control other factors which may correlate with

mortality and/or air pollution. It has been well established that health is related to income level, amount of education and perhaps other measures of social and economic status. For example, high income areas tend to have a greater concentration of medical personnel at their availability; densely populated areas may be more likely to have contagious diseases.

The U.S. Census Bureau generally provide measures of socioeconomic status aggregated on a citywide or community basis. Chicago is unique in that such information has been provided for sub-units of the city. In 1920 Chicago was divided into 76 subareas or community areas, "each of which was based on the assumption that local communities would have their own history, common interest, local business and organizations meeting their primary needs and that they would also be bounded by natural and artificial barriers".[29] Each area was fairly homogeneous ethnically and economically. Great changes have occurred in Chicago since 1920, so that these community areas no longer function as they were originally devised but they still remain a method of dividing the city into subareas in order to analyze change in the population, social, economic and residential structures.[29]

Based on 1970 U.S. population and housing census data and other information, the Council for Community Services in Chicago developed a set of indicators to measure income levels, environmental conditions, health status and social well-being in all communities.[29] Health status and social well-being indicators are considered to include mortality in their concept and indeed involve mortality statistics as factors to create these two indicators. Thus, only three indicators were used in the study: income levels, environmental conditions and educational levels. The following variables were used to develop the three indicators by the council:[29]

1. Adequate income and economic opportunity (Income Score)
   a. Median family income, 1970
   b. % of families receiving public aid, 1969
   c. % of white collar workers 16 years and over, 1970
   d. % of laborers and service workers 16 years and over, 1970
   e. % of unemployed persons age 16 years and over in civilian labor force, 1970
2. Basic material needs and optimal environmental conditions (Environmental Score)

a. % of year-round housing units lacking built-in heating facilities, 1970

b. % of occupied housing units lacking plumbing facilities, 1970

c. % of occupied housing units having more than one occupant per room, 1970

d. % of occupied housing units lacking an automobile, 1970

e. % of occupied housing units lacking an available telephone, 1970

f. Number of persons per square mile, 1970

g. Number of male juvenile deliquents committed to correctional institutions per 100 males ages 12-16, 1967-1972

h. Age-adjusted death rate from homicides in 1972

3. Adequate knowledge and skills (Education Score)

a. Median years of school completed for persons, 25 years of age, 1970

b. % of males 16-21 years of age, enrolled in school, 1970

c. % of persons, 25 years of age and over, completed high school, 1970

d. % of persons, 25 years of age and over, completed college, 1970

The council used a factor analysis, or specifically a principal components analysis, to develop index scores in 76 community areas for each indicator. In a principal components analysis, the first factor extracted accounts for the greatest proportion of the variance, the second factor, the second greatest proportion, etc. The factors in their analysis were not rotated. Loadings in the first factors in each goal area were taken as the weighting coefficients for that goal area; secondary factors, which accounted for less of the variance, were discarded. Their procedures are summarized as follows:

The variables to which the weighting coefficients apply were expressed in different units, and it was necessary to express them in comparable terms. This was done by transforming all raw scores to standard scores. It was desired that the low scores on any indicator should be zero, and the high score, 100. To accomplish this, a theoretical "best" score was calculated for each indicator by taking the "best" observed on each score variable associated with the indicator, multiplying each score by its weighting coefficient and summing the products. It should be emphasized that the "best" scores on the component variables were drawn from different communities; in no case did a single community have all the "best" scores related to a given community only. A similar procedure was used to generate a "worst" score on each indicator. Since the variables were expressed in the standard

score values, the range was considerably smaller than the 0-100 which was sought. One indicator, for example, had a low score of -11.5 and a high score of 13.5, making a range of 25. For this indicator, it was necessary to multiply each indicator score by 4 and to add 46 to each observation. This had the effect of putting its range in the desired basis without affecting the shape of the distribution of scores. A comparable procedure was followed for each of the other indicators; a theroretical "best" and "worst" score were determined; a multiplier was applied so that the "best score would be a 100 units greater than the "worst" score; and a constant was added to each observation so the low score would have a value of 0, and the high score, of 100.

(cited from reference 29, pp. 150-151)

Socioeconomic characteristics of ethnic populations can be reflected in the three indicators used in this study. In fact, the percentage of blacks in communities are highly correlated to income score (r=0.772) and environmental score (r=0.712). Also, correlations between percentages of blacks and age-adjusted death rates were close to those between age-adjusted death rates and one, two, or all of the three indicators. Accordingly, percentages of any ethnic populations were not included as control variables.

After computing correlation coefficients among dependent and independent variables, most correlations with income scores, environment scores and education scores indicated negative signs, so that original scores were subtracted from 100 to avoid negative signs in a correlation matrix. Thus, it is interpreted that a high score means low quality of environmental conditions, low income levels, and low education levels in communities.

Air Pollution Data

The city of Chicago Department of Environmental Control (DEC) monitors a variety of pollutants at measuring sites located throughout the city (Figure 1). This system, the Chicago Air Sampling Network (CASN), monitors such pollutants as TSP, O_3, NO_2, SO_2, and NO at up to 30 sites. An initial investigation of data available from the DEC pointed out that only TSP and SO_2 were monitored consistently in the period 1971-75.

Tapes were obtained from the EPA containing TSP and SO_2 measurements for the 5 year period. This tape consisted of data from the aforementioned CASN supplemented with measurements from 3 sites under jurisdiction of the State

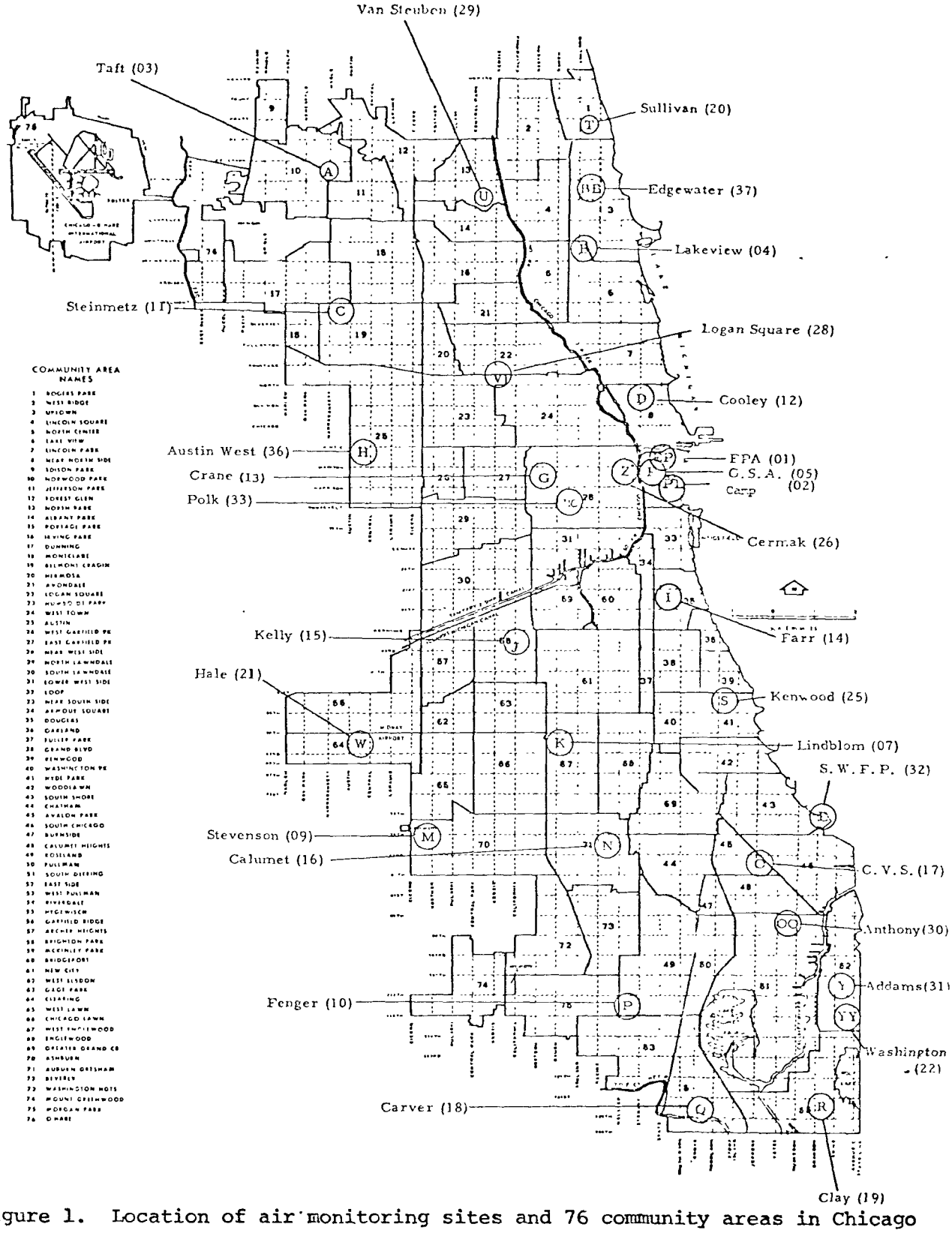

Figure 1. Location of air monitoring sites and 76 community areas in Chicago

Table 5.2 Yearly Averages of Sulfur Dioxide (part per million) in the Chicago Air Sampling Network for the Years 1971-75.

Site No.	Monitoring Site	1971 No. of obs.	1971 Average (ppm)	1972 No. of obs.	1972 Average (ppm)	1973 No. of obs.	1973 Average (ppm)
01	EPA	18	.0275	30	.0176	22	.0115
02	Camp	27	.0272	28	.0179	23	.0221
03	Taft	95	.0161	105	.0166	99	.0117
04	Lakeview	94	.0295	104	.0273	98	.0186
05	GSA	94	.0274	103	.0356	64	.0296
06	Austin	90	.0198	102	.0262	101	.0173
07	Lindblom	93	.0160	103	.0180	99	.0114
09	Stevenson	92	.0132	103	.0141	98	.0187
10	Fenger	95	.0166	104	.0224	99	.0161
11	Steinmetz	94	.0181	108	.0144	63	.0086
12	Cooley	93	.0304	105	.0250	65	.0222
13	Crane	89	.0190	103	.0150	57	.0134
15	Kelly	96	.0236	103	.0198	62	.0141
16	Calumet	95	.0151	113	.0221	63	.0160
17	Chgo. Voc.	91	.0073	94	.0143	65	.0139
18	Carver	87	.0241	103	.0223	62	.0160
19	Clay	92	.0217	106	.0188	64	.0144
20	Sullivan	91	.0188	104	.0219	63	.0160
21	Hale	95	.0166	105	.0273	63	.0207
22	Washington	93	.0328	105	.0274	65	.0206
25	Kenwood	66	.0274[a]	105	.0250	100	.0191
30	Anthony	36	.0231[a]	104	.0154	62	.0111
31	Adams	36	.0231[a]	105	.0199	65	.0263
32	SWFP	35	.0182[a]	104	.0159	64	.0155
Citywide Average			.0215		.0208		.0169

Note: a. Measurements are missing from January to July.
b. Measurements are missing in December.
c. Measurements are missing from October to November.
d. Measurements are missing from May to December in 1974 and from January to May in 1975, so that measurements for 1974 and 1975 were combined as one year's to calculate a 5-year average.
e. Yearly averages of 1974 and 1975 were estimated by measurements at site 15 (Kelly High School) which are highly correlated with those at site 13.
f. Measurements are missing from June to July.
g. Measurements are missing in August.

Table 5.2 con't Yearly Averages of Sulfur Dioxide (part per million) in the Chicago Air Sampling Network for the Years 1971-75

Site No.	Monitoring Site	1974		1975		5-year average
		No. of obs.	Average (ppm)	No. of obs.	Average (ppm)	(ppm)
01	EPA	13	.0144	20	.0158	.0174
02	Camp	24	.0142	26	.0189	.0201
03	Taft	61	.0102	60	.0069	.0123
04	Lakeview	48	.0099[c]	53	.0129	.0196
05	GSA	57	.0231	59	.0186	.0269
06	Austin	18	.0213[d]	34	.0102[d]	.0202
07	Lindblom	58	.0126	59	.0063	.0129
09	Stevenson	57	.0136	57	.0109	.0141
10	Fenger	61	.0144	58	.0113	.0162
11	Steinmetz	58	.0075	41	.0070[f]	.0111
12	Cooley	59	.0205	58	.0162	.0229
13	Crane	--	(.0091)[e]	--	(.0085)[e]	.0130
15	Kelly	58	.0114	56	.0106	.0159
16	Calumet	55	.0126	58	.0094	.0150
17	Chgo. Voc.	60	.0116	59	.0088	.0112
18	Carver	59	.0127	57	.0078	.0166
19	Clay	60	.0100	60	.0095	.0149
20	Sullivan	53	.0097	51	.0062	.0145
21	Hale	59	.0146	56	.0112	.0181
22	Washington	60	.0216	56	.0128	.0230
25	Kenwood	57	.0139	59	.0090	.0188
30	Anthony	57	.0033	38	.0068[g]	.0119
31	Adams	59	.0224	57	.0163	.0224
32	SWFP	55	.0078	44	.0057	.0126
Citywide Average			.0134		.0107	.0167

Note:
a. Measurements are missing from January to July.
b. Measurements are missing in December
c. Measurements are missing from October to November.
d. Measurements are missing from May to December in 1974 and from January to May in 1975, so that measurements for 1974 and 1975 were combined as one year's to calculate a 5-year average.
e. Yearly averages of 1974 and 1975 were estimated by measurements at site 15 (Kelly High School) which are highly correlated with those at site 13.
f. Measurements are missing from June to July.
g. Measurements are missing in August.

Table 5.3 Yearly Averages of Total Suspended Particulate ($\mu g/m^3$) in the Chicago Air Sampling Network for the years 1971-75

Site No.	Monitoring Site	1971 No. of obs.	1971 Average ($\mu g/m^3$)	1972 No. of obs.	1972 Average ($\mu g/m^3$)	1973 No. of obs.	1973 Average ($\mu g/m^3$)
01	EPA	25	115[a]	27	97[a]	26	88[a]
02	Camp	138	173	132	155	93	147
03	Taft	134	75	140	70	110	76
04	Lakeview	145	93	142	80	108	83
05	GSA	142	116	138	101	115	108
06	Austin	133	98	137	81	115	88
07	Lindblom	140	83	127	90	109	81
09	Stevenson	136	87	129	83	109	80
10	Fenger	139	93	137	79	113	79
11	Steinmetz	132	72	140	67	110	72
12	Cooley	136	131	125	116	114	126
13	Crane	135	121	133	102	98	104
14	Farr	140	109	135	87	96	79
15	Kelly	130	96	138	96	112	88
16	Calumet	135	91	132	79	110	82
17	Chgo. Voc.	139	98	138	84	114	82
18	Carver	133	106	136	101	54	92[e]
19	Clay	136	92	129	88	93	91
20	Sullivan	114	84	130	71	73	65
21	Hale	140	100	140	87	100	92
22	Washington	139	164	136	134	86	164
25	Kenwood	81	88[b]	139	80	112	76
28	Logan Square	127	103	135	85	92	77
29	Von Steuben	126	80	145	48	98	64
30	Anthony	73	99[c]	135	93	93	91
31	Adams	80	130	137	112	98	122
32	SWFP	75	92[d]	139	67	94	68
33	Polk	-	-	44	103[a]	53	127[a]
Citywide average			99.44		90.57		92.57

Note: a. Measured by U.S. EPA or Ill. EPA.
b. Measurements are missing from January to March.
c. Measurements are missing in January.
d. Measurements are missing in May.
e. Measurements are missing from September to December.
f. Measurements are missing from May to December.
g. Measurements are missing from January to May.
h. Measurements are missing from April to July.
i. Yearly averages of 1974 and 1975 were estimated by measurements at site 15 (Kelly High School) which are highly correlated with those at site 13.

Table 5.3 con't. Yearly Averages of Total Suspended Particulate ($\mu g/m^3$) in the Chicago Air Sampling Network for the years 1971-75.

Site No.	Monitoring Site	1974 No. of obs.	1974 Average ($\mu g/m^3$)	1975 No. of obs.	1975 Average ($\mu g/m^3$)	5-year average
01	EPA	7	80[a]	22	106[a]	97.2
02	Camp	22	120[a]	30	121[a]	143.2
03	Taft	122	75	106	60	71.2
04	Lakeview	92	74	88	70	80.0
05	GSA	109	103	101	95	104.6
06	Austin	33	82[f]	53	92[g]	88.9
07	Lindblom	101	60	94	63	75.4
09	Stevenson	111	70	98	75	79.0
10	Fenger	116	85	103	80	83.2
11	Steinmetz	115	65	65	67[h]	68.6
12	Cooley	105	114	98	94	116.2
13	Crane	-	(108)[i]	-	(102)[i]	107.4
14	Farr	90	84	73	77	87.2
15	Kelly	101	90	96	85	91.0
16	Calumet	110	81	102	69	80.4
17	Chgo. Voc.	110	91	104	75	86.0
18	Carver	82	73	86	73	89.0
19	Clay	93	88	96	78	87.4
20	Sullivan	85	63	83	57	68.0
21	Hale	83	80	92	68	85.4
22	Washington	84	153	98	148	152.6
25	Kenwood	98	70	94	65	75.8
28	Logan Square	90	81	79	72	83.8
29	Von Steuben	93	67	73	70	65.8
30	Anthony	96	95	83	86	92.8
31	Adams	97	132	97	105	120.2
32	SWFP	86	68	98	67	72.4
33	Polk	18	73[a]	59	90[a]	98.3
Citywide average			85.81		81.78	91.1

Note:
a. Measured by U.S. EPA or Ill. EPA.
b. Measurements are missing from January to March.
c. Measurements are missing in January
d. Measurements are missing in May.
e. Measurements are missing from September to December.
f. Measurements are missing from May to December.
g. Measurements are missing from January to May.
h. Measurements are missing from April to July.
i. Yearly averages of 1974 and 1975 were estimated by measurements at site 15 (Kelly High School) which are highly correlated with those at site 13.

of Illinois EPA. A listing of this tape was produced and the data was found to contain extensive error making it unreliable for use. Mimeographed copies of the same data were obtained from CASN which was used to correct errors in the original tapes. Several months were spent at a great cost, to produce a final, reliable air pollution file of SO_2 and TSP for the 5-year study period.

In 1971-73, TSP was monitored three times weekly and SO_2 was measured every 3rd day. Finally, in 1974-75, TSP continued to be measured as frequently as in 1973 but SO_2 was measured every 6th day. Thus, the frequency of measurements declined from about 115 per year to 95 per year at each site for TSP. SO_2 was monitored about 95 times per year in 1971 declining to about 55 per year in 1975. These numbers are lower than would be expected from the frequencies mentioned above, since every site was subject to random shutdowns of varying duration because of such factors as equipment failure. One site (Hyde Park) closed in 1971, and another (CRIB) was deleted because it was closed more than 50% of the time. The Crance site was closed during 1974-75, but its data was estimated by measurements taken at Kelly, a nearby site whose measurements were highly correlated with Crane's.

Because the frequency of measurements changed each year, it was decided to calculate yearly averages of SO_2 and TSP at each site, rather than a 5-year average based on combining all measurements. Such a procedure would have given more weight to data from the yearly part of the study period since pollutants were monitored more frequently then.

After 5 yearly averages were calculated at each site, a 5-year average was produced by taking a simple average of the 5 yearly averages. This was done for both SO_2 and TSP at each site (See Tables 5.2 and 5.3).

Estimation of Community Area Exposure Levels

Once 5-year averages for SO_2 and TSP levels were calculated at each monitoring site in the CASN system, a procedure was developed to estimate these pollutant levels for each community area (called CA exposure levels). These would be used as estimates of personal exposure for all residents of the particular CA.

Since not every CA had a monitoring site within its boundaries, a set of interpolation equations based on measurements at the sites were produced

Table 5.4

Estimated Community Exposure Levels to Total Suspended Particulate ($\mu g/m^3$) and Sulfur Dioxide (ppm) for Five Years and Interpolation Formulas Used for Each Community Area (C.A.)

C.A.	TSP ($\mu g/m^3$)	SO_2 (ppm)	Site No. used for TSP estimation	Site No. used for SO_2 estimation
1	68.0	.0145	20	20
2	66.9	.0145	½(20 + 29)	20
3	74.0	.0170	½(20 + 4)	½(20 + 4)
4	71.3	.0170	⅓(4 + 20 + 29)	½(4 + 20)
5	81.8	.0196	½(4 + 28)	4
6	80	.0196	4	4
7	93.3	.0213	⅓(4 + 12 + 28)	½ (4 + 12)
8	116.2	.0229	12	12
9	71.2	.0123	3	3
10	71.2	.0123	3	3
11	71.2	.0123	3	3
12	68.5	.0123	½(3 + 29)	3
13	65.8	.0134	29	½(3 + 20)
14	65.8	.0155	29	⅓(3 + 4 + 20)
15	68.5	.0117	⅓(3 + 11 + 29)	½(3 + 11)
16	71.5	.0143	⅓(4 + 11 + 29)	⅓(3 + 4 + 11)
17	68.6	.0111	11	11
18	68.6	.0111	11	11
19	68.6	.0111	11	11
20	67.2	.0111	½(11 + 28)	11
21	83.6	.0154	28	½(4 + 11)
22	83.6	.0213	28	½(4 + 12)
23	86.3	.0202	½(6 + 28)	6
24	99.9	.0180	½(12 + 28)	½(12 + 13)
25	88.9	.0202	6	6
26	98.2	.0166	½(6 + 13)	½(6 + 13)

Table 5.4 con't.

Estimated Community Exposure Levels to Total Suspended Particulate ($\mu g/m^3$) and Sulfur Dioxide (ppm) for Five Years and Interpolation Formulas Used for Each Community Area (C.A.)

C.A.	TSP ($\mu g/m^3$)	SO_2 (ppm)	Site No. used for TSP estimation	Site No. used for SO_2 estimation
27	98.2	.0166	½(6 + 13)	½(6 + 13)
28	102.9	.0130	½(13 + 33)	13
29	98.2	.0166	½(6 + 13)	½(6 + 13)
30	95.8	.0164	⅓(6 + 13 + 15)	⅓(6 + 13 + 15)
31	99.2	.0145	½(13 + 15)	½(13 + 15)
32	115	.0215	⅓(1 + 2 + 5)	⅓(1 + 2 + 5)
33	107	.0199	⅓(2 + 5 + 14)	⅓(2 + 5 + 32)
34	87.2	.0126	14	32
35	79.8	.0126	½(14 + 32)	32
36	74.1	.0157	½(25 + 32)	½(25 + 32)
37	78.5	.0157	⅓(14 + 32 + 25)	½(25 + 32)
38	78.5	.0157	⅓(14 + 32 + 25)	½(25 + 32)
39	75.8	.0188	25	25
40	75.6	.0159	½(7 + 25)	½(7 + 25)
41	75.8	.0188	25	25
42	75.8	.0188	25	25
43	80.9	.0150	½(17 + 25)	½(17 + 25)
44	83.2	.0131	½(16 + 17)	½(16 + 17)
45	86.0	.0112	17	17
46	89.4	.0116	½(17 + 30)	½(17 + 30)
47	86.0	.0112	17	17
48	86.0	.0112	17	17
49	83.2	.0162	10	10
50	88	.0141	½(10 + 30)	½(10 + 30)
51	89.7	.0145	⅓(18 + 19 + 30)	⅓(18 + 19 + 30)
52	136.4	.0227	½(22 + 31)	½(22 + 31)

Table 5.4 con't.

Estimated Community Exposure Levels to Total Suspended Particulate ($\mu g/m^3$) and Sulfur Dioxide (ppm) for Five Years and Interpolation Formulas Used for Each Community Area (C.A.)

C.A.	TSP ($\mu g/m^3$)	SO_2 (ppm)	Site No. used for TSP estimation	Site No. used for SO_2 estimation
53	86.1	.0164	½(10 + 18)	½(10 + 18)
54	89.0	.0166	18	18
55	87.4	.0149	19	19
56	85.4	.0181	21	21
57	88.2	.0170	½(21 + 15)	½(15 + 21)
58	91.0	.0159	15	15
59	91.0	.0159	15	15
60	89.1	.0143	½(14 + 15)	½(15 + 32)
61	83.2	.0144	½(7 + 15)	½(7 + 15)
62	88.2	.0170	½(15 + 21)	½(15 + 21)
63	83.2	.0144	½(7 + 15)	½(7 + 15)
64	85.4	.0181	21	21
65	82.2	.0161	½(9 + 21)	½(9 + 21)
66	77.2	.0135	½(7 + 9)	½(7 + 9)
67	75.4	.0129	7	7
68	77.9	.0140	½(7 + 16)	½(7 + 16)
69	80.4	.0150	16	16
70	79.0	.0141	9	9
71	80.4	.0150	16	16
72	80.9	.0151	⅓(9 + 10 + 16)	⅓(9 + 10 + 16)
73	81.8	.0156	½(10 + 16)	½(10 + 16)
74	81.1	.0152	½(9 + 10)	½(9 + 10)
75	83.2	.0162	10	10
76	69.9	.0117	½(3 + 11)	½(3 + 11)

utilizing the following rules:

1. If a CA contained a site within its boundary the site measurements was used as the exposure level.
2. If two or more sites were located in one CA, the exposure level was the average of measurements at all sites.
3. If a CA contained no stations, the two or three stations closest to the CA were identified. Then, the exposure level was calculated as the average of measurements at these sites.

Table 5.4 contains the exposure levels and the interpolation equations used to estimate them at each CA. Note that in some instances the SO_2 equations are not the same as those for TSP, caused by a number of stations monitoring TSP which did not monitor SO_2.

Multiple Regression Analysis

Multiple regression analysis was used, because not only the strength of a relationship between a dependent variable and a specific independent variable can be measured, holding other independent variable constant, but also the method provides us with a quantitative estimate of such a relationship. The dependent variables used were 5-year averages of age-adjusted death rates for 17 disease categories. Independent variables were (1) a 5-year average of community exposure levels of air pollution (TSP, SO_2 and TSPx SO_2), and (2) environmental, income and education scores in community areas. Means and standard deviations of these variables were shown in Table 5.5. Correlation coefficients between age-adjusted death rates and independent variables are listed in Table 5.6, in addition to a correlation matrix of independent variables in Table 5.7.

The general model to be tested in this study is expressed as follows:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n + \varepsilon$$

where Y = the age-adjusted death rate for a specific disease

β_0 = constant

$\beta_1 \beta_2 \ldots \beta_n$ = regression coefficients

$X_1, X_2 \ldots X_n$ = pollutant(s) and environmental and socioeconomic scores

n = the number of independent variables included in an equation

ε = random error term

Table 5.5 Means and Standard Deviations of Dependent and Independent Variables from 76 Community Areas in Chicago for the years 1971-75

Variables	Mean	Standard Deviation
Dependent Variables (age-adjusted death rates)		
(1) All causes excluding accidents, homicides and suicides	1018.64	204.03
(2) Malignant neoplasms	193.41	33.85
a. Neoplasms of digestive organs and peritoneum	61.37	14.66
b. Neoplasms of respiratory systems	42.84	9.98
c. Neoplasms of genito-urinary organs	32.61	9.42
(3) Heart disease	492.09	106.98
a. Ischemic heart disease	415.79	78.88
b. other heart diseases	76.29	55.93
(4) Cerebrovascular disease	91.86	21.25
(5) Arteriosclerosis	11.79	5.31
(6) Other circulatory diseases	19.05	7.12
(7) Diabetes mellitus	22.15	8.26
(8) Cirrhosis of liver	25.88	13.12
(9) Pneumonia and influenza	34.51	15.43
(10) Emphysema	7.95	3.33
(11) Congential anomalies and diseases of early infancy	29.72	10.20
(12) All others not included in (2) - (11) and excluded accidents, homicides and suicides	90.22	31.03
	(number of deaths per 100,000 persons)	
Independent Variables		
(1) Total suspended particulate ($\mu g/m^3$)	83.44	12.52
(2) Sulfur dioxide (ppm)	0.0155	0.0030
(3) TSP x SO_2	1.3137	0.4216
(4) Environment score	26.67	17.67
(5) Income score	40.47	18.77
(6) Education score	61.78	12.48

Table 5.6

Correlation Coefficient Between Age-Adjusted Death Rates by Major Causes and Independent Variables

	TSP	SO_2	TSP x SO_2	Income	Environment	Education
1. All non-accidental causes	0.325**	0.152	0.243*	0.531**	0.581**	0.399**
2. Malignant neoplasms	0.022	-0.064	-0.035	0.251	0.144	0.107
a. digestive organs and peritoneum	0.078	-0.054	-0.007	0.325**	0.166	0.201
b. respiratory systems	0.206	0.102	0.154	0.267*	0.262*	0.204
c. genito-urinary organs	-0.016	-0.028	-0.034	0.306**	0.206	0.048
3. Heart Disease	0.358**	0.150	0.268*	0.373**	0.446**	0.380**
a. ischemic heart disease	0.240*	0.040	0.156	0.002	0.052	0.175
b. other heart diseases	0.346**	0.231*	0.293*	0.710**	0.781**	0.480**
4. Cerebrovascular disease	0.015	0.065	0.046	0.378**	0.396**	0.084
5. Arteriosclerosis	0.026	-0.007	0.032	-0.374**	-0.170	-0.234*
6. Other Circulatory disease	0.190	0.153	0.187	0.455**	0.480**	0.241*
7. Diabetes mellitus	0.330**	0.108	0.206	0.556**	0.498**	0.562**
8. Cirrhosis of liver	0.466**	0.241*	0.363**	0.497**	0.646**	0.536**
9. Pneumonia and Influenza	0.262*	0.273*	0.279*	0.715**	0.768**	0.396**

Table 5.6 con't

Correlation Coefficient Between Age-Adjusted Death Rates by Major causes and Independent Variables

		TSP	SO_2	TSPxSO_2	Income	Environment	Education
10.	Emphysema	0.374**	0.157	0.288*	-0.120	-0.062	0.100
11.	Congenital anomalies and disease of early infancy	0.161	0.106	0.108	0.507**	0.555**	0.249*
12.	Other non-accidental causes	0.309**	0.158	0.220	0.531**	0.581**	0.400**

Note: High scores of income, education and environment mean low income levels, low education levels and low quality of environmental conditions in the community, respectively. Computations was based on 76 community areas.

Asterisks on coefficients indicate levels of significance; * $\alpha = .05$ and ** $\alpha = .01$

Table 5.7

Correlation Matrix of Independent Variables

	TSP	SO_2	TSP x SO_2	Income	Environment	Education
TSP	1.000	0.559**	0.861**	0.244*	0.308**	0.350**
SO_2	0.559**	1.000	0.891**	0.138	0.305**	0.116
TSP x SO_2	0.861**	0.891**	1.000	0.169	0.307**	0.218
Income	0.244	0.138	0.169	1.000	0.871**	0.699**
Environment	0.308**	0.305**	0.307**	0.871**	1.000	0.619**
Education	0.350**	0.116	0.218	0.699**	0.619**	1.000

Note: High scores of income, education, and environment mean low income levels, low education levels and low quality of environmental conditions in the community, respectively.

Computation was based on 76 community areas.

Asterisks on coefficients indicates levels of significance: * α = .05 and ** α = .01

In computation procedures, the air pollution variable(s) was first included into the equation. As a second step, environmental, economic and education scores were introduced by stepwise methods. Four models were created for each disease category according to type(s) of the pollutant(s) as follows:

1. total suspended particulate (TSP) alone
2. sulfur dioxide (SO_2) alone
3. SO_2 and TSP together
4. TSP x SO_2, or an interaction between TSP and SO_2 alone

In a final list of the models, we chose only those which met the criteria that (1) a coefficient of the pollutant is significant or close to the significant level at $p<.10$, and (2) a most meaningful model is chosen if none of four models has a significant coefficient of the pollutant.

DAILY ANALYSIS

Death Statistics

As described in the previous section on cross-sectional analysis, daily mortality information in a computer tape was arranged according to 76 community areas, 11 age groups and 17 causes. If the number of daily deaths from a specific disease category is too small, it is not appropriate for multiple regression analysis. For example, the mean number of daily deaths from respiratory diseases (ICD No. 470-474, 480-486, 492) was about 4 deaths per day which is too small for this type of analysis. Hence, the number of city-wide daily deaths was obtained for the period 1971-75 from two disease categories only:

1. total deaths excluding homicides, suicides and accidents
2. deaths from all heart diseases
   (See ICD code number in Table 5.1)

Daily deaths from all community areas were combined to obtain one city-wide total for each of the two disease categories for each day over the years 1971-75.

Climatological Data

The U.S. Department of Commerce, National Oceanic and Atmospheric

Administration (NOAA), maintains a local climatological data bank comprising Chicago area stations. The data bank currently monitors meteorological variables at 3 stations within the city of Chicago: on the University of Chicago campus, at Midway Airport, and near Buckingham Fountain in the Loop. However, the Midway Airport site presented the most concise set of daily data maintained by a complete and precise monitoring system. For this purpose, daily climatological variables in the model are represented by Midway Airport (as published by U.S. Department of Commerce, NOAA).

The following variables were included in the study:

1. Maximum temperature (^{o}F)
2. Minimum temperature (^{o}F)
3. Average temperature (^{o}F)
4. Amount of precipitation (inches)
5. Amount of snowfall (inches)
6. Average wind speed (miles per hour)
7. Percent relative humidity (%)
8. Percent of possible hours sunshine (%)
9. Proportion of sky covered by clouds (measured in tenths).

Aerometric Data

Air contaminant data was obtained from the CASN described earlier. This consisted of measurements for TSP and SO_2 at 27 and 23 sites, respectively. The frequency of measurements changed during the 5 year period for which data was received. TSP was measured three times weekly and SO_2 was measured twice weekly during 1971-72. In 1973 TSP and SO_2 was measured every 3rd day. Finally for the years 1974-75, TSP continued to be measured every 3rd day but SO_2 was measured every 6th day.

For each day that a pollutant was measured, a citywide pollution average was calculated by combining measurements at all sites of the network. Usually, not all sites were open on a given day. To alleviate any error in citywide averages which could be caused by missing data a TSP average was deleted from the study if less than 20 sites were included in its calculation, and a SO_2 average was deleted if less than 18 sites were used.

Multiple Regression Analysis

To develop a quantitative relationship between daily mortality and air pollution, multiple regression analysis was chosen as the best method available. A variant of the stepwise approach was used.

The dependent variable of interest was number of deaths due to:

1. All causes excluding homicide, suicide and accident
2. All heart diseases

The general approach was to force the pollutant(s) of interest into the regression equation on the first step and then to enter in the most significant of the climatological variables one at a time on succeeding steps. For both death categories three models were developed:

1. TSP as the pollutant of interest
2. SO_2 as the pollutant of interest
3. The interaction between TSP and SO_2 as the pollutant (calculated by multiplying the daily average of TSP by that of SO_2)

Initially all 9 climatological variables were included as possible variables. Later both minimum and maximum daily temperature were excluded, since average daily temperature was found to be more highly correlated with the disease categories, and it would be more representative of meteorological conditions for the day as a whole.

Day of the week correction: Since both mortality and air pollution levels previously have been found to vary according to the day of the week, this was taken into account for the model development. Day of the week being a categorical rather than a quantitative variable, a set of 6 dummy (0-1) variables was forced into each regression equation. They were coded as follows: Let the set of dummy variables be represented by the vector (M, T, W, Th, F, S); then Sunday is coded as (0,0,0,0,0,0), Monday as (1,0,0,0,0,0), Tuesday as (0,1,0,0,0,0), etc. Since Sunday is represented by 6 zeros then the regression coefficients for the remaining days must be interpreted as the additional number of deaths expected for that day of the week compared to the average for all Sundays. Emphasis should be placed on the total effects of the day-of-week variables rather than singling out one or some of them.

The statistical significance of the day-of-week variables added to the

regression equation was tested by calculating the F ratio of the difference between the two R^2's before and after adding the day-of-week variables as follows:

$$F = \frac{(R^2_{y.12\cdots k_1} - R^2_{y.12\cdots k_2})/(K_1-K_2)}{(1-R^2_{y.12\cdots k_1})/(N-K_1-1)}$$

where $R^2_{y.12\cdots k_1}$ = the squared multiple correlation coefficient for the regression of Y (number of deaths in this study) on K_1 variables (the larger coefficient) after adding the day-of-week variables (K_1=7 in this study); and $R^2_{y.12\cdots k_2}$ = the squared multiple correlation coefficient for the regression of Y on the K_2 variables, where K_2= the number of independent variables before adding the day-of-week variables (K_2=1 in this study because only one pollutant of interest was included before adding six day-of-week variables to the equation). This F value was shown as total effects of the day-of-week variables on the last row of Tables 6.3 and 6.4 in DAILY ANALYSIS of section 6.

Lag-day effects: The models considered so far are those for which pollution levels are taken for the same day as death occurred (0 day lag). There is also the possibility that pollutants may continue to affect mortality totals for a number of days after. So two additional series of models were developed to account for any possible lag effects.

The first set of models used a three day lag. In other words death total and climatological data were matched to air contaminant levels three days before. All models described above were recalculated using this three day lag. Similarly, a six day lag effect was tested.

SECTION 6

RESULTS AND DISCUSSION

CROSS-SECTIONAL ANALYSIS

Multiple regression equations in cross-sectional analysis are summarized in Table 6.1. Two models were chosen in the age-adjusted death rate for all causes excluding accidents, homicides and suicides (called death rate I). Model I-1 included total suspended particulate (TSP) and the environment score because including the income score and the education score did not improve the regression. The F ratio to test the overall goodness of fit of model I-1 was significant; $20.64 > F(.01,2,60)^{*} = 4.98$. The F value of TSP, 2.68, was not significant but close to the significant level, $F(.10,1,60) = 2.79$. Thirty-six per cent of the variation of death rate I was accounted for by model I-1 including TSP and the environment score. Adding SO_2 to model I-1 resulted in increasing the F values of both TSP and the environment score in model I-2. The F value of TSP, 4.35, was in fact significant ($>F(.05,1,60)=4.00$). SO_2 itself was inversely correlated with the death rate, but its F value was not significant. Wherever SO_2 was included in the regression equation, its regression coefficient was negative except in model X - 2 (emphysema). Accordingly, the effect of TSP on death rate I was estimated by using model I-1 which did not include SO_2 as an independent variable. Table 6.2 shows the results of a percentage decrease in the age-adjusted rate when a 25 percent reduction in total suspended particulate or sulfur dioxide was introduced in selected models, controlling for the environment, the income, and the education scores. It is estimated that a 25% reduction in TSP (20.86 $\mu g/m^3$) would decrease the age-adjusted death rate of total deaths (non-accidental deaths) by 5.36% (54.65 deaths per 100,000) in Chicago.

---

*$F(.01,2,60)=4.98$ is obtained from the F-distributions table. In parentheses, a first figure indicates a significant level to be tested, and the second and third figures show degrees of freedom which are equal or closest to (but smaller than) degrees of freedom in the study sample.

TABLE 6.1
MORTALITY MODELS IN CROSS-SECTIONAL ANALYSIS

		I. All causes excluding accidents, homicides & suicides		II. Malignant neoplasms All neoplasms	a. Digestive organs and peritoneum	b. Respiratory systems	c. Genito-urinary organs	III. Heart disease All heart diseases
		(1)	(2)					
Multiple R		.60	.61	.31	.41	.30	.39	.50
R^2		.36	.38	.10	.17	.09	.15	.25
F ratio		20.64***	14.47***	1.51	3.54**	3.72**	3.17**	12.38***
d.f.		2,73	3,72	5,70	4,71	2,73	4,75	2,73
Constant		636.1	684.7	190.2	45.5	27.9	39.8	258.2
TSP	reg.	2.62	3.88**	.15	.09	.12	.003	2.08**
	s.e.	1.60	1.86	.40	.16	.09	.106	.91
	F	(2.68)	(4.35)	(.15)	(.32)	(1.71)	(.00)	(5.27)
SO_2	reg.		-10312	-908	-258		-217	
	s.e.		7861	1678	686		424	
	F		(1.72)	(.29)	(.14)		(.26)	
$TSP \times SO_2$	reg.							
	s.e.							
	F							
Environment	reg.	6.14***	6.39***	-.50	-.39*			2.25***
	F	(29.24)	(31.11)	(1.09)	(3.98)			(12.21)
Income	reg.			1.04**	.57***	.12**	.27***	
	F			(4.80)	(9.97)	(4.00)	(12.38)	
Education	reg.			-.39			-.24**	
	F			(.75)			(4.09)	

Note: a. For each independent variable, regression coefficient, standard error (for pollutant only) and F statistic are represented on lines of reg., s.e. and F, respectively.

b. The level of significance is marked as asterisk: *(α=.10), **(α=.05), and ***(α=.01)

Table 6.1 con't

		III. Heart disease a. Ischemic heart disease	b. Other heart disease (1)	b. Other heart disease (2)	IV. Cerebro-vascular disease	V. Arterio-sclerosis	VI. Other circulatory disease
Multiple R		.31	.79	.80	.48	.49	.50
R^2		.10	.63	.64	.23	.24	.25
F ratio		2.52*	30.29***	24.39***	4.16***	7.69***	6.06***
d.f.		3,72	4,71	5,70	5,70	3,72	4,71
Constant		244.1	-26.3	-14.2	109.3	15.3	13.69
TSP	reg.	1.28*	.62*	.83**	-.63	.03	.05
	s.e.	.75	.35	.42	.23	.05	.06
	F	(2.90)	(3.13)	(4.02)	(.08)	(.33)	(.61)
SO_2	reg.			-1682	-149		
	s.e.			1762	973		
	F			(.91)	(.02)		
TSPxSO_2	reg.						
	s.e.						
	F						
Environment	reg.		1.96***	2.12***	.39	.19***	.13
	F		(17.11)	(17.79)	(1.94)	(8.42)	(2.26)
Income	reg.	-.98	.60	.50	.40	-.26***	.11
	F	(2.23)	(1.53)	(1.03)	(2.09)	(19.67)	(1.58)
Education	reg.	1.69	-.42	-.48	-.59**		-.11
	F	(2.71)	(.80)	(1.04)	(5.04)		(1.58)

Table 6.1 con't

		VII. Diabetes mellitus		VIII. Cirrhosis of liver		IX. Pneumonia and influenza	X. Emphysema		
		(1)	(2)	(1)	(2)		(1)	(2)	(3)
Multiple R		.62	.63	.74	.75	.79	.45	.33	.40
R^2		.39	.39	.55	.56	.63	.21	.11	.16
F ratio		15.24***	11.48***	21.30***	17.99***	29.79***	6.42***	2.93**	4.52***
d.f.		3,72	4,71	4,71	5,70	4,71	3,72	3,72	3,72
Constant		-3.6	-2.4	-15.6	-10.2	19.21	-1.62	1.97	2.72
TSP	reg.	.10	.13*	.25***	.34***	.10	.10***		
	s.e.	.06	.08	.09	.11	.10	.03		
	F	(2.40)	(2.82)	(7.32)	(10.19)	(1.04)	(11.95)		
SO_2	reg.		-225		-744			192	
	s.e.		315		453			126	
	F		(.51)		(2.70)			(2.31)	
TSPxSO_2	reg.								2.26**
	s.e.								.87
	F								(6.69)
Environment	reg.			.58***	.65***	.51***			
	F			(21.97)	(25.20)	(14.80)			
Income	reg.	.14**	.14**	-.30**	-.34***	.29**	-.07**	-.07**	-.07**
	F	(6.11)	(6.36)	(5.56)	(7.14)	(4.77)	(6.38)	(6.24)	(6.31)
Education	reg.	.19**	.18**	.28**	.25**	-.30**	.06	.09**	.81*
	F	(4.62)	(4.07)	(5.41)	(4.39)	(5.26)	(2.16)	(5.13)	(3.93)

Table 6.1 con't

		XI. Congenital anomalies and diseases of early infancy	XII. All others excluding accidents, homicides and suicides	
			(1)	(2)
Multiple R		.58	.83	.85
R^2		.33	.70	.71
F ratio		8.94***	40.72***	35.06***
d.f.		4,71	4,71	5,70
Constant		26.0	45.4	58.5
TSP	reg.	.03	.28	.51**
	s.e.	.09	.18	.20
	F	(.12)	(2.57)	(6.31)
SO_2	reg.			-1826**
	s,e.			865
	F			(4.46)
$TSPxSO_2$	reg.			
	s.e.			
	F			
Environment	reg.	.27**	.95***	1.12***
	F	(5.33)	(15.97)	(20.79)
Income	reg.	.14	.68***	.58**
	F	(1.33)	(7.74)	(5.56)
Education	reg.	-.19	-.51**	-.58**
	F	(2.65)	(4.75)	(6.30)

TABLE 6.2

ESTIMATED EFFECT ON AGE-ADJUSTED DEATH RATE OF A 25 PERCENT REDUCTION IN TOTAL SUSPENDED PARTICULATE OR SULFUR DIOXIDE

Model		Type of pollutant	% decrease in age-adjusted death rate
I-1	All causes excluding accidents, homicides and suicides	TSP	5.36 (54.65)
III	All heart diseases	TSP	8.82 (43.39)
III-a	Ischemic heart disease	TSP	6.42 (26.70)
III-b-1	Other heart disease	TSP	16.95 (12.93)
VII-1	Diabetes mellitus	TSP	9.39 (2.08)
VIII-1	Cirrhosis of liver	TSP	20.13 (5.21)
X-1	Emphysema	TSP	26.16 (2.08)
X-2	Emphysema	SO_2	9.35 (0.74)
XII-1	All other causes excluding accidents, homicides and suicides	TSP	6.47 (5.84)

Note: 1) Each figure in parentheses after a % decrease in the age-adjusted death rate indicates the decreased number of deaths per 100,000 persons in Chicago.

2) A 25 percent reduction based on the 5-year average of the community exposure levels (TSP: 83.44 $\mu g/m^3$, SO_2: 0.0155 ppm) is equivalent to 20.86 $\mu g/m^3$ of TSP and 0.003875 ppm of SO_2.

This estimate is almost twice as much as the estimate made by Lave and Seskin[30] who associated a 50% reduction in total suspended particulate and sulfate with a 4.7% to 6.3% decrease in the crude death rate of all causes (including accidental deaths) and with a 4.8% to 5.5% decrease in the age-sex-race adjusted death rate of all causes (including accidental deaths). The difference in the estimation of the mortality decrease between the present study and Lave-Seskin's study is explained by the following reasons: (1) their study used the death rate of all causes (our study excluded accidents, homicides and suicides which are considered to be unrelated to air pollution), (2) their study used 69 to 117 SMSAs across the nation and the annual averages of particulate and sulfate in one SMSA might be far from the average exposure levels of the population in that area (our study used 76 Chicago community areas of which each is geographically much smaller than a SMSA), (3) our study did not include sulfate measurements and our estimates were based on TSP only in model I-1, and (4) their study controlled for the population density, the percentage of the SMSA population aged sixty-five and older, the percentage of the nonwhite population, the percentage of the families with incomes below the poverty level and the logarithm of the SMSA population. (Our study controlled for the environment score, the income score and the education score which represented the environmental condition, the income level and the education level in the community area and accounted for most of the control factors listed in Lave-Seskin's study). Considering these comparisons, our model might more appropriately reflect air pollution effects on mortality than Lave-Seskin's model.

Deaths from all malignant neoplasms (called death rate II) had an R^2 of 0.10 (Table 6.1). Neither TSP nor SO_2 was significant. Only the income score had a significant association with death rate II. The positive coefficient means that there is a tendency for cancer death rates to be higher in the poor rather than the rich communities in Chicago. This tendency was the same in three major subgroups of cancer. None of these subgroups showed significant coefficients of the air pollutants in their models. In model II-b (respiratory systems), a standard error of TSP, 0.09, was smaller than its regression coefficient of 0.12, but its F value was not significant ($1.71<F(.10,1,60)=2.79$). In Lave-Seskin's study cited above, their model for the cancer death rate of respiratory systems indicated that only the mean

level of sulfate pollution approached significance and no indices of particulate pollution had meaningful impact on the death rate. Also, their model for the cancer death rate from digestive organs showed a significant coefficient of the minimum levels of sulfate pollution but an insignificant coefficient of particulate pollution. Deaths from cancer of the genitourinary organs could not be compared because Lave and Seskin did not analyze for this subgroup. The common finding between our study and theirs is the insignificant association of particulate pollution with the death rate from malignant neoplasms, controlling for socioeconomic factors. Accordingly, no attempt was made to estimate the cancer mortality change which might be associated with changes in air pollution.

For the death rate for heart disease, all four models showed significant coefficients of TSP (Table 6.1). The model for all heart diseases had an R^2 of 0.25 with a significant F ratio ($p<.01$) and a significant coefficient of the environment score ($p<.01$) and TSP ($p<.05$). The income score and the education score were not included in the model because of their meaningless contributions to the regression. It is expected that communities with deteriorated environmental conditions and high concentrations of TSP would have a high death rate attributable to heart diseases. Lave-Seskin's models of cardiovascular disesses indicated that the minimum sulfate level was statistically significant, while the maximum suspended particulate measure approached statistical significance. The model of ischemic heart disease in Table 6.1 had an R^2 of 0.10 with a significant F ratio ($p<.10$). Only TSP reached the significant level ($p<.10$) although both the education score and the income score approached significance. Two models were listed in other heart disease because the addition of SO_2 to model III-b-1 increased the F values of both TSP and the environment score in model III-b-2, although a coefficient of SO_2 itself was negative and insignificant. Model III-b-2 confirmed the significant association of TSP with the age-adjusted death rate for other heart disease. It is interesting that the model for other heart disease had a greater R^2 than for ischemic heart disease (0.63 vs. 0.10). Environmental conditions in the communities had the greatest impact on other heart disease among the independent variables, while environmental conditions in the communities would be less important in ischemic heart disease than income levels and education levels. According to model III-a,

it is expected that the communities with higher particulate levels, higher income levels and lower education levels would have a higher death rate for ischemic heart disease. As shown in Table 6.2, a 25% reduction in TSP (20.86 $\mu g/m^3$) would decrease the age-adjusted death rates for all heart diseases by 8.82% (43.39 deaths per 100,000 persons), for ischemic heart disease by 6.42% (26.70 deaths per 100,000 persons) and for other heart disease by 16.95% (12.93 deaths per 100,000 persons). Although the percentage decrease in the death rate for ischemic heart disease was smaller than for other heart disease, the frequency of deaths from ischemic heart disease (415.79 deaths per 100,000 persons as the average of 76 community areas) was much higher than from other heart disease (76.29 deaths per 100,000 persons). One way to examine the accuracy of the estimation for all heart diseases is to compare the decreased portion of the death rate for all heart disease (43.39 deaths per 100,000 persons) against the sum of the decreased portion of the death rates for ischemic heart disease and for other heart disease (26.70 + 12.93 = 39.63 deaths per 100,000 persons). The difference of 3.76 deaths per 100,000 persons might be caused by the differences in the types of independent variables included in three models (III, III-1, and III-b-1).

Model IV (cerebrovascular disease) had an R^2 of 0.23 with a significant F ratio, $4.16 > F(.01,5,60) = 3.34$ (Table 6.1). Both TSP and SO_2 were not significant. The education score was negatively and significantly ($p<.05$) associated with the death rate. This implies that the communities with more educated people had a tendency to have a higher death rate for cerebrovascular disease than those with less educated people. The income and the environment scores were close to the significant level at $p<.10$.

Model V (arteriosclerosis) had an R^2 of 0.24 and an F ratio of 7.69 ($>F(.01,3,60) = 4.43$) with the inclusion of TSP (insignificant), the income score (significant, $p < .01$) and the environment score (significant, $p < .01$). The communities with high income levels and/or poor environmental conditions might be expected to have a higher death rate for arteriosclerosis than the other communities.

The analysis for other circulatory disease did not indicate any significant and meaningful contributions of the pollution measures to the regression model. No independent variables were significant in Model VI although an F

ratio of 6.06 to test the overall goodness of fit of this regression was significant (Table 6.1).

For diabetes mellitus, two models (VII-1 and VII-2) were chosen (Table 6.1). Model VII-1 included TSP with an F value of 2.40 ($<F(.10,1,60) = 2.79$), in addition to the income score and the education score which were both significant ($p<.05$). The addition of SO_2 to the first model increased the F value of TSP which reached the significant level ($2.82 > F(.10,1,60)=2.79$), although the coefficient of SO_2 was negative and insignificant. Accordingly, TSP is considered to be meaningfully associated with the death rate for diabetes mellitus. Because of the negative coefficient of SO_2, model VII-1 was used to estimate the amount of the mortality rate decrease in diabetes mellitus by introducing a 25% reduction in TSP (20.86 $\mu g/m^3$). The estimated annual decrease in the death rate was 9.39% which was equivalent to 2.08 deaths per 100,000 persons by diabetes mellitus as an underlying cause of death in Chicago. Because diabetes actually affects the cardiovascular and renal systems, diabetics living in polluted communities might have a higher risk of having heart disease and/or kidney failure as a secondary cause of death than the others. Both models VII-1 and VII-2 imply that the communities with more low-income residents and/or more uneducated residents would have a higher death rate for diabetes mellitus than the others; that is, diabetics in low socioeconomic communities might not be receiving proper medical treatments and dietary control.

Model VIII-1 (cirrhosis of the liver) had an R^2 of 0.55 with an F ratio of 21.30 ($>F(.01,4,60)=3.65$), as shown in Table 6.1. All independent variables included in the model were strongly related to the age-adjusted death rate for cirrhosis of the liver. TSP showed the second strongest association with the death rate following the environmental score; the F value of TSP was 7.32 which was significant at the .01 level ($>F(.01,1,60)=7.08$). The addition of SO_2 to the model increased the F value of TSP to 10.19, but the coefficient of SO_2 was negative again. Model VIII-1 and VIII-2 indicate that communities with deteriorated environmental conditions, high economic levels and/or low education levels in addition to high particulate levels would have a higher death rate for cirrhosis of the liver. Based on model VIII-1 (Table 6.2), a 25% reduction (20.86 $\mu g/m^3$) of TSP from its average of 76 community areas, might result in a 20.13% annual decrease (5.21 deaths

per 100,000 persons) in the age-adjusted death rate for cirrhosis of the liver. The significant relationship between total suspended particulate and the death rate for cirrhosis of the liver in the study is consistent with the study results from Winkelstein and Gay[31] who analyzed the death rates among whites according to five economic levels (median family income) and four TSP levels. Their contingency table revealed a strong inverse association between economic level and cirrhosis mortality and a similar, strong but positive association between cirrhosis mortality and suspended particulate air pollution, although they did not take into account educational and environmental conditions which were significantly associated with cirrhosis mortality in our study.

Model IX (pneumonia and influenza) is one of the best prediction equations developed because the variation of the death rates explained by the model was 63% with a highly significant F ratio, $29.79 > F(.01,4,60)=3.65$ (Table 6.1). Three scores (environment, income and education) were all significant, but TSP was not significant. The model shows that environmental conditions in the community would be the most critical factors among the independent variables. If there is any air pollution effects on this disease group, the acute effects of air pollution might be more serious than its chronic effects considering the nature of the disease.

Model X-1 (Emphysema) indicated a strong and positive relationship between TSP and the age-adjusted death rate for this disease. The regression coefficient of TSP was over three times its standard error with an F value of 11.95 ($>F(.01,1,60) = 7.08$). The income score reached a significant level and the education score approached significance. The R^2 was 0.21 with a significant F ratio at the .01 level. Model X-2 revealed, for the first time throughout the analyses, a meaningful and positive association between SO_2 and emphysema mortality, although an F value of SO_2 was not significant but close to the significant level ($2.31 < F(.10,1,60)=2.79$). Model X-3 showed a significant relationship between emphysema mortality and a product of TSP and SO_2, which had an F value of $6.69 > F(.05,1,60) = 4.00$. The negative coefficient of the income score and the positive coefficient of the education score imply that communities with high income levels and/or low education levels tended to have a higher death rate for emphysema than the

other communities. Based on model X-1 (Table 6.2), a 25% reduction (20.86 $\mu g/m^3$) of TSP from its average in the 76 community areas would decrease the annual age-adjusted death rate for emphysema by 26.16% (2.08 deaths per 100,000 persons). Also using model X-2, a 25% reduction (0.0039 ppm) of SO_2 would decrease the death rate for emphysema by 9.35% (7.44 deaths per one million persons).

Model XI (congenital anomalies and diseases of early infancy) had an R^2 of 0.33 with a significant F ratio of 8.94 at the .01 level (Table 6.1). Among the independent variables included in the model, only the environment score was significant and the education score was close to the significant level. TSP was not significant at all. The positive association of the environment score with this disease group means that a higher death rate for the diseases in early infancy would be expected in the communities with a deteriorated environment than with a better environment.

Model XII-1 or XII-2 (all other causes excluding accidents, homicides and suicides) had the greatest R^2, 0.70 or 0.71, with a significant F ratio ($p < .01$). The most strongly associated with the death rate for this disease group was the environment score (significant, $p < .01$), the income score (significant, $p < .01$), the education score (significant, $p < .05$) and TSP (insignificant, close to the .10 level; $2.57 < F(.10,1,60) = 2.79$) in order. Again, the addition of SO_2 to the model resulted in a great increase in the F value of TSP (significant at the .05 level), but a coefficient of SO_2 was negative and significant at the .05 level. Because there is a concensus among researchers that high levels of SO_2 would be harmful but would not improve the health conditions of a population, the negative and significant coefficient of SO_2 appeared to be caused by high correlations between SO_2 and TSP as well as other environmental and socioeconomic variables. Such multicollinearity is one of the greatest problems in using a multivariate regression method. To avoid the correlation between air pollution variables (SO_2 and TSP in this study) a regression equation was created to include each pollutant separately with other independent variables. In this study the relationship between SO_2 and TSP is not considered to be causal; the level of SO_2 would not affect that of TSP, and vice versa. Therefore, we did not use the equation which included both SO_2 and TSP as independent variables for the actual estimation of a mortality

change. In regard to the model for all other causes excluding accidents, homicides and suicides, model XII-1 was used to estimate a percent change of its death rate after a 25% reduction of TSP occurs in the city. As shown in Table 6.2, the death rate for this disease group would decline by 6.47 percent (5.84 deaths per 100,000 persons).

One way to examine the reliability of the mortality changes due to a reduction in TSP is to compare a portion of decrease in the death rate for all causes without accidental deaths (54.65 deaths per 100,000 persons) against the sum of portions of decrease in the death rates for ischemic heart disease, other heart disease, diabetes mellitus, cirrhosis of the liver, emphysema and all other causes without accidental deaths (26.70 + 12.93 + 2.08 + 5.21 + 2.08 + 5.84 = 54.84 deaths per 100,000 persons). The difference in the death rate change between the first group and the sum of the others was only 0.19 deaths per 100,000 persons. Consequently, the estimated mortality changes in Table 6.2 are regarded as reliable and logical estimates, controlling for the environmental and socioeconomic conditions in the community areas in Chicago.

DAILY ANALYSIS

All Causes Except Accidents, Homicides And Suicides (Non-Accidental Deaths)

A total of 9 models were created to examine acute effects of total suspended particulate and sulfur dioxide on deaths from all causes excluding accidents, homicides and suicides (non-accidental deaths only), as shown in Table 6.3. Three regression models are presented in each of three sets of anlyses (A. day of onset, B. three day lag and C. six day lag) according to the inclusion of pollutant(s) in the equation. For example, models I-0, II-0, and III-0 included TSP, SO_2 and $TSPxSO_2$, respectively, in the regression equation for the day of death onset. In the same way, models for heart disease deaths are presented in Table 6.4.

As shown in Table 6.3, the models for the day of onset had R^2's of 0.14 (I-0), 0.21 (II-0) and 0.19 (III-0) with significant F ratios ($p<.01$) and significant coefficients of all pollution indices ($p<.01$). Model I-0 shows that total suspended particulate was significantly related to daily non-accidental deaths ($p<.01$), even holding constant daily average temperature, precipitation, wind spped and humidity as well as the day-of-week variables

TABLE 6.3
MORTALITY MODELS IN DAILY ANALYSIS: ALL CAUSES EXCEPT ACCIDENTS, HOMICIDES AND SUICIDES

	A. Day of onset			B. Three day lag		
	I-O	II-O	III-O	I-3	II-3	III-3
Multiple R	.38	.46	.44	.36	.42	.39
R^2	.14	.21	.19	.13	.18	.15
F ratio	8.02**	9.42**	8.86**	7.10**	9.72	7.70**
d.f.	11,531	10,351	9,339	11,531	8,353	8,340
Constant	90.44	88.47	100.1	91.08	96.80	103.82
TSP	.079**			.014		
	(19.95)			(.67)		
SO_2		386.5**			330.8**	
		(23.75)			(20.15)	
TSPxSO_2			1.71**			1.36**
			(13.76)			(8.29)
Average Temp.	-.19**	-.08	-.17**	-.19**	-.11**	-.17**
	(44.10)	(3.04)	(21.25)	(37.79)	(6.96)	(21.63)
Precipitation	3.64*	7.48**	8.93**			
	(4.27)	(10.80)	(15.32)			
Snowfall						
Windspeed	.30	.35		.28		
	(3.42)	(3.30)		(3.21)		
Hunidity	.07			.13*		
	(2.59)			(6.02)		
Sunshine				.03		
				(2.57)		
Skycover						
Day-of-week variables (coefficents only)						
Monday	-.97	1.51	1.11	-2.06	.25	-.77
Tuesday	-.33	1.17	1.82	-1.30	.78	.10
Wednesday	-5.11	-.07	-.99	-.50	-.59	-1.44
Thursday	-2.69	-.58	-.32	-5.56	-4.29	-5.22
Friday	-8.86	-4.08	-4.65	2.29	3.17	2.78
Saturday	-5.39	-4.49	-5.01	-6.52	-2.58	-3.57
Total effects[3] (F value)	3.52**	1.28	1.43	4.09**	1.60	1.69

Note: 1) For each independent variable, the first figure indicates its regression coefficient and the second figure in parentheses shows its F value

2) The level of significance is marked as asterisk: *(α = .05) and **(α = .01).

3) Calculation of this F value is shown in Multiple Regression Analysis of Daily Analysis in Section 5.

TABLE 6.3 con't

MORTALITY MODELS IN DAILY ANALYSIS: ALL CAUSES EXCEPT ACCIDENTS, HOMICIDES AND SUICIDES

	C. Six day lag		
	I-6	II-6	III-6
Multiple R	.38	.47	.46
R^2	.14	.22	.21
F ratio	7.90**	9.20**	7.52**
d.f.	11,530	11,349	12,335
Constant	83.32	90.72	89.91
TSP	.034		
	(3.62)		
SO_2		321.8**	
		(17.62)	
TSPxSO_2			1.36**
			(7.67)
Average Temp.	-.19**	-.13**	-.19**
	(35.73)	(8.16)	(20.74)
Precipitation		5.14*	5.25*
		(4.09)	(4.18)
Snowfall			
Windspeed	.49**	.29	.29
	(9.03)	(2.20)	(1.98)
Humidity	.13*		
	(5.44)		
Sunshine	.04*	.03	.08*
	(4.21)	(2.25)	(4.71)
Skycover			.64
			(2.67)
Day-of-week variables			
Monday	6.51	5.97	6.22
Tuesday	.58	-.85	-1.93
Wednesday	2.97	1.54	1.55
Thursday	-3.98	-2.79	-2.93
Friday	-.35	-2.78	-4.27
Saturday	2.37	-5.71	5.62
Total effects[3] (F value)	4.37**	3.73**	4.37**

Note: 1) For each independent variable, the first figure indicates its regression coefficient and the second figure in parentheses shows its F value.

2) The level of significance is marked as asterisk: *(α=.05) and **(α=.01).

3) Calculation of this F value is shown in Multiple Regression Analysis of Daily Analysis in Section 5.

which were the significant addition to the model according to an F value of 3.52 ($p<.01$). In model II-0, coefficients of SO_2 and precipitation were significant at $p<.01$, but both coefficients of average temperature and wind speed were not significant at $p<.05$ although they were meaningful in the model. In model III-0 which examined effects of an interaction between TSP and SO_2 on daily non-accidental deaths, coefficients of the pollutant ($TSPxSO_2$), average temperature and precipitation were all significant at $p<.01$.

All models for the three-day lag had a smaller R^2 than those for the day of onset (0.13 in model I-3 vs. 0.14 in model I-0, 0.18 in II-3 vs. 0.21 in II-0, and 0.15 in III-3 vs. 0.19 in III-0), and showed significant F ratios at $p<.01$. In model I-3, a coefficient of TSP was not significant at all, but a coefficient of average temperature remained significant at $p<.01$, while humidity became significant ($p<.05$) with meaningful inclusion of wind speed and sunshine. The day-of-week variables had singificant impact on daily non-accidental deaths as a whole according to an F value of 4.09 ($p<.01$) in model I-3. In model II-3, a coefficient of SO_2 (330.8) decreased slightly, compared to that in model II-0 (386.5), but remained highly significant ($p<.01$). Only average temperature was included in model II-3 as a climatological index, while model II-0 had average temperature, precipitation and wind speed. Model III-3 included both the pollution variable ($TSPxSO_2$) and average temperature with significant F values ($p<.01$).

The models for the six-day lag had an increase in R^2's with significant F ratios ($p<.01$), as compared to those for the three-day lag; 0.14 (I-6) vs. 0.13 (I-3), 0.22 (II-6) vs. 0.18 (II-3), and 0.21 (III-6) vs. 0.15 (III-3). The strength of the relationship between the pollution index and non-accidental deaths in the six-day lag analysis was close to the one in the three-day lag analysis, except model I-6 in which TSP approached significance at $p<.05$. Model I-6 included four climatological variables of which all F values were significant; temperature and wind speed at $p<.01$, and humidity and sunshine at $p<.05$. In model II-6, a coefficient of SO_2, 321.8, was smaller than either 386.5 in model II-0 or 330.8 in model II-3, but its F value, 17.62, was significant at $p<.01$ and greatest among the independent variables included in model II-6. In model III-6, a coefficient of $TSPxSO_2$, 1.36, was the same as that in model III-3, although five

climatological variables (average temperature, precipitation, wind speed, sunshine and sky cover) were included in model III-6, as compared to the inclusion of average temperature only as a climatological variable in model III-3. In regard to the day-of-week variables, F values to test total effects were significant in all three models for the six day lag ($p<.01$).

Based on the models described above, both sulfur dioxide and suspended particulate could account for some portion of the variation of daily non-accidental deaths on the day when such deaths occurred, even holding weather and day-of-week variables constant. It is also implied that sulfur dioxide levels on the third and/or the sixth days prior to death occurence could have serious impact on daily non-accidental deaths. In addition, a significant coefficient of $TSPxSO_2$ shows the implication that sulfur dioxide and particulate might have synergetic effects on mortality from non-accidental causes on the day of death occurence, the third day and the sixth day prior to death occurence.

Based on models I-0, II-0 and III-0, a percentage of decrease in the number of daily non-accidental deaths related to a 25% reduction in air pollution can be estimated by using the averages of daily non-accidental deaths and air pollution levels (TSP, SO_2 and $TSPxSO_2$). The following average figures were used; 98.86 $\mu g/m^3$ (TSP), 0.018 ppm (SO_2) and 1.9345 ($TSPxSO_2$) which were corresponded to 95.47 deaths, 95.81 deaths and 95.77 deaths, respectively. Portion of a 25% reduction was 24.715 $\mu g/m^3$ (TSP), 0.0045 ppm (SO_2) and 0.4836 ($TSPxSO_2$). A product of each pollution decrease and a regression coefficient is portion of decrease in deaths; 2.045% (1.95 deaths)-TSP, 1.815% (1.74 deaths)-SO_2 and 0.867% (0.83 deaths)-$TSPxSO_2$.

Lave and Seskin[30] created linear models to examine the relationship between daily mortality and daily air pollution in Chicago by using total deaths (including accidental deaths), climatological variables (mean temperature, rainfall and wind speed) and air pollution measurements (SO_2, NO_2, NO and H_2C) from the Continuous Air Monitoring Program (CAMP) during the 3-year period (September 1962 - May 1964). Their model included air pollution levels on the day of death occurence and on five preceding days as independent variables in the same equation, as well as with climatological variables and day-of-week variables. They estimated that a 50 percent reduction in air pollution (as measured by sulfur dioxide) was

associated with a 5.4 percent reduction in daily deaths. In other words, a 25 percent reduction in SO_2 would decrease daily deaths by 2.7 percent, as compared to our estimate, a 1.815 percent decrease in daily non-accidental deaths. The difference in daily mortality decreases between their model and ours may be caused by the differences in the dependent variables (total deaths included accidental deaths vs. total non-accidental deaths), SO_2 measurements (one measurement from CAMP vs. a city-wide average based on measurements from more than 20 monitoring sites) and inclusion of SO_2 measurements (on five preceding days vs. on the day of death onset only).

Heart Disease

Daily mortality models for heart disease are summarized in Table 6.4. In the analysis for the day of death onset, the strength of the relationship between daily deaths from heart disease and daily air pollution levels was less strong than between daily total non-accidental deaths and daily air pollution levels according to regression coefficients and F values of air pollutants. R^2's were 0.13 in model I-0, 0.17 in model II-0 and 0.16 in model III-0 with all significant F ratios ($p<.01$). Regression coefficients of the pollutants were 0.038 (TSP), 174.8 (SO_2) and 0.90 ($TSPxSO_2$) with all significant F values ($p<.01$). Among the climatological variables included in the models, daily average temperature had a strong association with daily mortality from heart disease in all three models. Precipitation was included in models II-0 and III-0 and wind speed was in model I-0 only. The addition of the day-of-week variables to the model appeared significant ($p<.01$) in model I-0 only.

Models for the three-day lag had little changes in R^2's from those for the day of onset. Both coefficient and F values decreased considerably; especially, a coefficient of TSP became insignificant in model I-3, while coefficients of both SO_2 and $TSPxSO_2$ remained significant at $p<.01$ and at $p<.05$, respectively. Daily average temperature on the third day prior to death occurence had a significant association with daily mortality from heart disease. The day-of-week variables significantly contributed to model I-3 according to an F value of 3.78 ($p<.01$).

All three models for the six-day lag had an increase in an R^2; 0.15 (model I-6), 0.22 (II-6) and 0.22 (III-6) with all significant F ratios

TABLE 6.4
MORTALITY MODELS IN DAILY ANALYSIS: HEART DISEASE

	A. Day of onset			B. Three day lag		
	I-0	II-0	III-0	I-3	II-3	III-3
Multiple R	.36	.41	.40	.35	.42	.40
R^2	.13	.17	.16	.12	.17	.16
F ratio	8.72**	8.06**	7.21**	9.09**	8.29**	7.18**
d.f.	9,533	9,352	9,339	8,534	9,352	9,339
Constant	47.13	47.83	50.68	51.23	46.83	50.36
TSP	.038**			.012		
	(11.16)			(1.25)		
SO_2		174.8*			156.6**	
		(12.77)			(11.75)	
TSPxSO_2			.90**			.62*
			(8.64)			(4.42)
Average Temp.	-.12**	-.08**	-.11**	-.12**	-.07**	-.11**
	(42.76)	(9.18)	(22.27)	(48.05)	(8.10)	(20.61)
Precipitation		3.46*	3.88*			
		(5.40)	(6.57)			
Snowfall					1.54	1.48
					(3.25)	(2.87)
Wind Speed	.19					
	(3.37)					
Humidity						
Sunshine						
Skycover						
Day-of-week variables (coefficients only)						
Monday	1.67	1.64	1.58	-2.79	-1.05	-1.49
Tuesday	.41	.03	.21	-.46	.98	.94
Wednesday	-2.38	-1.24	-1.45	-.39	-.78	-1.16
Thursday	-1.14	-1.38	-1.20	-2.38	-1.44	-1.85
Friday	-5.96	-4.00	-4.20	1.71	2.15	2.03
Saturday	-3.14	-3.62	-3.89	-3.89	-1.75	-2.18
Total effects[3] (F value)	3.74**	1.83	1.87	3.78**	1.63	1.75

Note: 1) For each independent variable, the first figure indicates its regression coefficient and the second figure in parentheses shows its F value.
2) The level of significance is marked as asterisk: *(α=.05) and **(α=.01).
3) Calculation of this F value is shown in Multiple Regression Analysis of Daily Analysis in Section 5.

TABLE 6.4 con't
MORTALITY MODELS IN DAILY ANALYSIS: HEART DISEASE

	C. Six day lag I-6	II-6	III-6
Multiple R	.38	.47	.47
R^2	.15	.22	.22
F ratio	8.24**	10.00**	8.57**
d.f.	11,530	10,350	11,336
Constant	43.13	46.00	45.39
TSP	.014		
	(1.62)		
SO_2		132.1**	
		(7.73)	
TSP×SO_2			.49
			(2.62)
Average Temp.	-.13**	-.10**	-.13**
	(42.81)	(14.60)	(28.35)
Precipitation			
Snowfall			
Windspeed	.26*	.26*	.24
	(6.31)	(4.61)	(3.66)
Humidity	.055		
	(2.40)		
Sunshine	.024	.019	.045*
	(3.65)	(2.38)	(3.85)
Skycover			.35
			(2.04)
Day-of-week variables (coefficient only)			
Monday	3.55	3.42	3.48
Tuesday	-.29	-1.26	-1.80
Wednesday	.31	-.16	-.34
Thursday	-3.41	-3.18	-3.23
Friday	-1.94	-3.18	-3.98
Saturday	.30	-3.99	-3.97
Total effects[3] (F value)	4.85*	4.77**	5.33**

Note: 1) For each independent variable, the first figure indicates its regression coefficient and the second figure in parentheses shows its F value.
2) The level of significance is marked as asterisk: *(α=.05) and **(α=.01).
3) Calculation of this F value is shown in Multiple Regression Analysis of Daily Analysis in Section 5.

($p<.01$). Only a coefficient of SO_2 remained significant at $p<.01$, while a coefficient of $TSPxSO_2$ lost significance and a coefficient of TSP was insignificant. Daily average temperature remained highly significant in all three models. Both wind speed and sunshine appeared in all three models as significant or meaningful contributions to them. All F values to test total effects of the day-of-week variables were significant at the .01 level.

Overall, a significant coefficient of SO_2 in models II-0, II-3 and II-6 implies that high sulfur dioxide levels on the day of death onset, the third and the sixth days before death would increase the number of daily deaths from heart disease. Also, a significant coefficient of TSP in model I-0 implies that high levels of particulate matter on the day of death occurence would increase the number of daily deaths from heart disease. In addition, a significant coefficient of $TSPxSO_2$ in model III-0 and III-3 has the implication that high levels of particulate matter and/or sulfur dioxide on the day of death onset and on the third day before death occurence might have synergetic effects on an increase in the number of daily deaths from heart disease.

A percentage of decrease in the number of daily deaths from heart disease can be estimated when a 25% reduction in air pollution (TSP, SO_2 and $TSPxSO_2$) is applied to models I-0, II-0 and III-0 in Table 6.4. Average figures used for calculation were 98.86 $\mu g/m^3$ (TSP) and 45.86 heart disease deaths in model I-0, 0.018 ppm (SO_2) and 46.28 heart disease deaths in model II-0, and 1.9345 ($TSPxSO_2$) and 46.28 heart disease deaths in model III-0. Portion of a 25% reduction was 24.715 $\mu g/m^3$ (TSP), 0.0045 ppm (SO_2) and 0.4836 ($TSPxSO_2$). A product of a regression coefficient and decreased portion of air pollution is portion of decrease in deaths; 2.048% (0.94 deaths)-TSP, 1.717% (0.79 deaths)-SO_2, and 0.940% (0.435 deaths)-$TSPxSO_2$.

COMMENT

Factors which may affect mortality are (1) demographic (age, sex, race, urban-rural and migration), (2) socioeconomic (income, education, environmental living conditions and occupation), (3) personal (nutrition, medical care, smoking habits, exercise habits and genetic factors), and (4) environmental (air quality or air pollution, water quality, occupational exposure to toxic substances, and climate). An ideal study on the association between air pollution and mortality would control for all the factors listed above. However, such a study is not feasible at present because much of the data is not available, especially on personal factors, lacking a survey to follow up the family of the deceased. (Even with a survey, such past information is often quite limited and unreliable). Nonetheless, there is an urgent need to measure the strength of the relationship between air pollution and mortality; at present this can be accomplished only by using existing and limited information.

In our study the multiple regression method was used to quantitate the association between air pollution and mortality, controlling for other related factors to mortality and/or air pollution. It is not appropriate to include all available factors in a regression equation, not only because F statistics to test an overall goodness of fit of the regression and to test significance of each regression coefficient are lowered if a factor added to the equation does not have any meaningful and significant association with a dependent variable (or mortality), but also because each addition of an independent variable to the equation loses one degree of freedom.

Accordingly, the factors to be included in the regression analysis should be selective. Such factors are considered to affect mortality, and to be independent of each other. In our cross-sectional analysis, three socioeconomic indices which were developed by the Council for Community Services were applied because they are considered to be the most comprehensive and representative to reflect the socioeconomic status of each community area. Although the three indices (income, environment and education) were interrelated to each other (see Table 5.7), they were all included in the analysis because the meaning and characteristic of each index is different from the other.

In regard to four groups of factors which may influence mortality, it is important to identify which factors are not controlled for and to discuss whether or not such factors were critical to determine the association between air pollution and mortality.

Among demographic factors, sex, race and migration were not controlled in our study, while death rates were adjusted for ages and an urban-rural factor was not concerned because only the urban population was used. Concerning the sex factor, the sex composition of our study population was assumed to have minimum effects on mortality of males and females combined (more than 80% of sex ratio out of 76 community areas fell between 0.85 and 0.95 which are considered to be in a normal range). To examine if air pollution affects male's mortality differently from female's, it is recommended that mortality models be created by sex. Among different racial and ethnic groups, blacks have a notably high mortality rate. The percentage of blacks in community areas were highly correlated with both environment and income scores (see page 18). Therefore, it is considered that socioeconomic characteristics of blacks were well reflected by both environment and income scores. Thus, the percentage of blacks was not included in our final analysis. It is assumed that socioeconomic and health-related characteristics of other ethnic groups in the community would be reflected on the three socioeconomic scores used in the study. Another factor which was not controlled is migration. The percentage of age-specific population changes at the community level is not known for the study year 1971-75. An increase or decrease in the community population during the study period might create an under- or an over-estimation of a death rate. Also, a shift in the age composition would cause a biased estimate of a death rate. We used the 1970 census population as the denominator to calculate the average age-adjusted death rate for the years 1971-75, based on the assumption that the population change for that period would be negligible and would not bias the overall results. One way to check migration effects on mortality in the future is to compare the age-adjusted death rate from our study with those based on the mid-year population between 1970 and 1980, when the U.S. population and housing census will be carried out.

Regarding socioeconomic factors, three factors (income, education and environmental conditions) were discussed earlier. The occupational factor was

partly reflected on the income score of which the development process contained the percentage of white collar workers 16 years and over, and the percentage of laborers and service workers 16 years and over from the 1970 population census. Probably more precise occupational information would be from death certificates. Unfortunately the occupational code was deleted in computer tapes obtained from the State of Illinois Department of Public Health. Persons in certain types of occupation could be exposed to toxic chemicals which may be more critical to cause illness than ambient air pollution.

The third factor, or the personal factor, was not controlled in our study because no information was available. Nutrition and medical care would be reflected on the income level, while there is no evidence that the percentage of persons having smoking havits, exercise habits and genetic weakness are different from one community to the other.

The fourth factor, the environmental factor, was a major concern in our study. Water quality was regarded as the same throughout the city of Chicago which supplies water from Lake Michigan. Also, climate was considered to be constant in the cross-sectional analysis. Total suspended particulate and sulfur dioxide measurements were used as air pollution indices because they were the most reliable and consistent measurements covering the entire city area during the study period. There is no question about the need to measure occupational exposure in the future, but the lack of information forced to exclude this factor in the present study.

Despite the exclusion of some control factors discussed above, it should be emphasized that the positive association between air pollution and mortality in our study cannot be denied, since the most significant factor to influence mortality, or the socioeconomic factor, was included in our models. The strong association between socioeconomic status and mortality was supported by Nagi and Stockwell[32]. Their study showed that the percentage of excess deaths due to major leading causes increases as socioeconomic status decreases. Furthermore, a 1975 study by William Kruvant and co-workers[33] showed the tendency that more persons in low socioeconomic status were clustered in heavily polluted areas than those in high socioeconomic status. Accordingly, it is essential to control for the socioeconomic factor to examine the association between air pollution and mortality. Failure to

control for the socioeconomic factor might create an apparently high relationship between air pollution and mortality.

It is more important to control climatological and day-of-week factors in daily analysis than the others in the cross-sectional analysis. Since the number of days included in the analysis was sufficiently large enough to obtain a stable estimate, it is quite likely that air pollution does have acute effects on mortality, based on our significant findings.

REFERENCES

1. Winklestein, W., Jr., S. Kanter, E.W. Davis, C.S. Maneri and W.E. Mosher. 1967. "The Relationship of Air Pollution and Economic Status to Total Mortality and Selected Respiratory System in Men, pt. I: Suspended Particulates," Archives of Environmental Health Vol. 14, p. 162.

2. Zeidberg, L.D., R.J.M. Horton, and E. Landau. 1967. "The Nashville Air Pollution Study, pt. V: Mortality from Diseases of the Respiratory System in Relation to Air Pollution," Archives of Environmental Health Vol. 15, p. 214.

3. Lepper, M.H., N. Shioura, B. Carnow, S. Andelman, and L. Lehrer. 1969. "Respiratory Diseases in an Urban Environment," Industrial Medicine Vol. 38, p. 36.

4. Mills, C.A. 1943. "Urban Air Pollution and Respiratory Diseases," American Journal of Hygiene Vol.37, p. 131.

5. Mills, C.A. 1952. "Air Pollution and Community Health," American Journal of Medical Sciences Vol. 224, p. 403.

6. Stocks, P. 1959. "Cancer and Bronchitis Mortality in Relation to Atmospheric Deposit and Smoke," British Medical Journal Vol. 1, p. 74.

7. Stocks, P. 1960. "On the Relations Between Atmospheric Pollution in Urban and Rural Localities and Mortality from Cancer Bronchitis and Pneumonia, with Particular Reference to 3:4 Benzopyrene, Beryllium, Molybdenum, Vanadian and Arsenic," British Journal of Cancer Vol. 14, p. 397.

8. Ashley, D.J.R. 1969. "The Distribution of Lung Cancer and Bronchitis in England and Wales," British Journal of Cancer Vol. 21, p. 187.

9. Ashley, D.J.R. 1969. "The Environmental Factors in the Aetiology of Gastric Cancer," British Journal of Preventive and Social Medicine Vol. 23, p. 187.

10. Gardner, M.J., M.D. Crawford, and J.N. Morris. 1969. "Patterns of Mortality in Middle and Early Old Age in the County Boroughs of England and Wales," British Journal of Preventive and Social Medicine Vol. 23, p. 133.

11. Buck, S.F., and D.A. Brown. 1964. Mortality from Lung Cancer and Bronchitis in Relation to Smoke and Sulfur Dioxide Concentration, Population Density, and Social Index (London, Tobacco Research Council) Research Paper No. 7.

12. Wicken, A.J., and S.F. Buck. 1964. Report on a Study of Environmental

Factors Associated with Lung Cancer and Bronchitis Mortality in Areas of North East England (London, Tobacco Research Council) Research Paper No. 8

13. Buck, S.F. and A.J. Wieken. 1967. "Models for Use in Investigating the Risk of Mortality from Lung Cancer and Bronchitis," Applied Statistics Vol. 16, p. 185.

14. Daly, C. 1959. "Air Pollution and Causes of Death," British Journal of Preventive and Social Medicine Vol. 13, p. 14.

15. Collins, J.J., H.S. Kasap, and W.W. Holland, 1971. "Environmental Factors in Child Mortality in England and Wales," American Journal of Epidemiology Vol. 93, p. 10.

16. Toyama, T. 1964. "Air Pollution and Health Effects in Japan," Archives of Environmental Health Vol. 8, p. 153.

17. Ziedberg, L.D., R.J.M. Horton and E. Landau. 1967. "The Nashville Air Pollution Study, pt. IV: Cardiovascular Disease Mortality in Relation to Air Pollution," Archives of Environmental Health Vol. 15, p. 225.

18. Enterline, P.E., A.E. Rikli, M.J. Sauer, and M. Hyman. 1960. "Death Rates for Coronary Heart Disease in Metropolitan and Other Areas," Public Health Reports Vol. 75, p. 759.

19. Sauer, H.T., G.H. Payne, C.R. Council and J.C. Terrell. 1966. "Cardiovascular Disease Mortality Patterns in Georgia and North Carolina," Public Health Reports Vol. 81, p. 455.

20. Friedman, G.D. 1967. "Cigarette Smoking and Graphic Variation in Coronary Heart Disease Mortality in the United States," Journal of Chronic Diseases Vol. 20, p. 769.

21. Gardner, M.J., M.D. Crawford, and J.N. Morris. 1969. "Patterns of Mortality in Middle and Early Old Age in the County Boroughs of England and Wales," British Journal of Preventive and Social Medicine Vol. 23, p. 133.

22. Greenburg, L., M.B. Jacobs, B.M. Drolette, F. Field and M.M. Braverman. 1962. "Report of an Air Pollution Incident in New York City, November 1953," Conference Report, Public Health Report Vol. 77, p. 7.

23. Greenburg, L., M.B. Jacobs, C.L. Erhardt, M. Glasser and J.I. Reed. 1967. "Air Pollution, Influenza and Mortality in New York City," Archives of Environmental Health Vol. 15, p. 430.

24. Gore, A.T., and C.W. Shaddick. 1958. "Atmospheric Pollution and Mortality in the County of London," British Journal of Preventive and Social Medicine Vol. 12, p. 104.

25. Glasser, M., and L. Greenburg, and F. Field. 1967. "Mortality and Morbidity During a Period of High Levels of Air Pollution," Archives of Environmental Health Vol. 15, p. 684.

26. Glasser, M., and L. Greenburg. 1971. "Air Pollution, Mortality and Weather," Archives of Environmental Health Vol. 22, p. 334.

27. Schimmel, H., and L. Greenburg. 1972. "A Study of the Relation of Pollution to Mortality," Journal of the Air Pollution Control Association Vol. 22, p. 607.

28. U.S. Bureau of the Census: Census of Population and Housing: 1970 Census Tracts, Final Report PHC (1)-43, Chicago, IL. SMSA.

29. Coucnil for Community Services in Metropolitan Chicago: 1975. Community Analysis Project, Report No. 1: Chicago Problems Analysis, Publication No. 7509.

30. Lave, L.B., and Seskin, E.P.: 1977. Air Pollution and Human Health. The Johns Hopkins University Press, Baltimore.

31. Winkelstein, W., and Gay, M.L.. 1971. "Suspended Particulate Air Pollution-Relationship to Mortality from Cirrhosis of the Liver." Arch. Environ. Health, Vol. 22, No. 1, pp. 174-177.

32. Nagi, M.H., and E.G. Stockwell. 1973. "Socioeconomic Differentials in Mortality by Cause of Death," Health Services Reports, Vol. 88, No. 5, pp. 449-456.

33. McCaull, J. 1976. "Discriminatory Air Pollution - If poor, don't breathe," Environment, Vol. 18, No. 2, pp.26-31.

TECHNICAL REPORT DATA
(Please read Instructions on the reverse before completing)

1. REPORT NO.	2.	3. RECIPIENT'S ACCESSION NO.
EPA-600/1-79-034		

4. TITLE AND SUBTITLE
Model for measuring the health impact from changing levels of ambient air pollution :Mortality study

5. REPORT DATE
August 1979

6. PERFORMING ORGANIZATION CODE

7. AUTHOR(S)
Tsukasa Namekata, Bertram W. Carnow, Domenic J. Reda, Eileen O"Farrell, and James R. Marselle

8. PERFORMING ORGANIZATION REPORT NO.

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Occupational and Environmental Medicine Program
School of Public Health, at the
University of Illinois, P.O. Box 6998
Chicago, IL 60680

10. PROGRAM ELEMENT NO.
1AA816

11. CONTRACT/GRANT NO.
68-02-2492

12. SPONSORING AGENCY NAME AND ADDRESS
Health Effects Research Lab RTP, NC
Office of Research & Development
U.S. EPA
Research Triangle Park, North Carolina 27711

13. TYPE OF REPORT AND PERIOD COVERED
Final Report

14. SPONSORING AGENCY CODE
EPA/600/11

15. SUPPLEMENTARY NOTES

16. ABSTRACT

The major objective of this study is to answer the questions," Is a recent mortality decline in the city of Chicago caused by a decrease in the amount of major air pollutants such as suspended particulate and sulfur dioxide?"

Based on multiple regression analysis for the cross-sectional analysis, a percentage decrease in the age-adjusted death rates was estimated when a 25 percent reduction in TSP in Chicago for the period 1970-75, was applied to the models developed. The age-adjusted death rate for non-accidental causes would be decreased by 5.36% (54.65 deaths per 100,000 persons) in Chicago. A percentage decrease in the death rates by cause was estimated to be 8.82% (all heart diseases), 6.42% (ischemic heart disease), 16.95% (other heart disease), 9.39% (diabetes mellitus), 20.13% (cirrhosis of the liver), 26.16% (emphysema) and 6.47% (other non-accidental causes).

Models developed in daily analysis imply that there would be possible acute effects of daily air pollution concentrations (both SO_2 and TSP, in addition to their interaction) on daily mortality changes (both all non-accidental causes and heart diseases), controlling for weather and day-of-week effects.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group
Air Pollution, Mortality Health Effects, Regression Analysis		06F

18. DISTRIBUTION STATEMENT	19. SECURITY CLASS *(This Report)*	21. NO. OF PAGES
PUBLIC	Unclassified	74
	20. SECURITY CLASS *(This page)*	22. PRICE
	Unclassified	

EPA Form 2220-1 (Rev. 4-77) PREVIOUS EDITION IS OBSOLETE