

**LINKING ENVIRONMENTAL EXPOSURES AND HEALTH EFFECTS USING
EXISTING DATA TO EXPLORE THE RELATIONSHIPS BETWEEN ENVIRONMENT
AND CHRONIC DISEASES**

by

Xiaohui Xu

B. Med., Zhejiang University, 1999

M. Med., Fudan University, 2002

Submitted to the Graduate Faculty of
Graduate School of Public Health in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

University of Pittsburgh

2007

UNIVERSITY OF PITTSBURGH

Graduate School of Public Health

This dissertation was presented

by

Xiaohui Xu

It was defended on

June 6th, 2007

and approved by

Dissertation Advisor:

Evelyn O. Talbott Dr. P.H., Professor
Department of Epidemiology,
Graduate School of Public Health
University of Pittsburgh

Committee member:

Vincent C. Arena Ph.D., Associate Professor
Department of Biostatistics, Graduate School of Public Health,
University of Pittsburgh

Committee member:

Devra Lee Davis Ph.D., Visiting Professor,
Department of Epidemiology, Graduate School of Public Health,
University of Pittsburgh

Committee member:

Ravi K. Sharma Ph.D., Assistant Professor,
Department of Behavioral Community Health Sciences,
Graduate School of Public Health, University of Pittsburgh

Committee member:

Conrad D Volz Dr. P.H., Assistant Professor,
Department of Environmental and Occupational Health,
Graduate School of Public Health, University of Pittsburgh

Copyright © by Xiaohui Xu

2007

**LINKING ENVIRONMENTAL EXPOSURES AND HEALTH EFFECTS USING
EXISTING DATA TO EXPLORE THE RELATIONSHIPS BETWEEN
ENVIRONMENT AND CHRONIC DISEASES**

Xiaohui Xu, PhD

University of Pittsburgh, 2007

The environment plays an important role in the health of communities. However, few health systems exist at the state and/or local levels to efficiently track the potential health effects associated with environmental exposure. The objectives of this dissertation are 1) to use secondary data for assessing the possible associations between health outcomes and environmental exposure and/or hazard; 2) to explore possible methods of data linkage and analyses which can be used by state and local environmental health tracking agencies and 3) to bring positive contributions to the development of national Environmental Public Health Tracking Network (EPHT). In this project, the Three Mile Island (TMI) cohort data (1979-1995) and Pennsylvania (PA) Cancer registry data were used to evaluate the associations between cigarette smoking and adult leukemia. A case-crossover analysis was performed with PA cardiopulmonary hospital admission data and local air pollution data to assess the health effects of air pollutants on cardiopulmonary disease before and after the elimination of a major point source of air pollution. A case-control study was also conducted to examine the associations between term low birth weight and particulate air pollution. The results showed that cigarette smoking could increase the risk of acute myeloid leukemia (AML). In addition, particulate air pollution is significantly associated with cardiovascular hospitalization and low birth weight in term infant.

In conclusion, the findings suggest that environmental hazards have adverse health effects on a number of health endpoints. Secondary data can be a great resource for environmental public health tracking, which is of public health relevance. The use of existing data is an effective way to assess the potential health effects associated with environmental exposures after an appropriate study design with a feasible data linkage and correct methods of data analyses was developed.

Keyword: Environmental exposure, Secondary data, Chronic disease

TABLE OF CONTENTS

PREFACE.....	XII
1.0 INTRODUCTION.....	1
1.1 ENVIRONMENTAL EXPOSURES AND CHRONIC DISEASES.....	1
1.2 NATIONAL ENVIRONMENTAL PUBLIC HEALTH TRACKING NETWORK.....	3
1.2.1 THE CURRENT STATUS OF ENVIRONMENTAL HEALTH TRACKING SYSTEMS.....	3
1.2.2 THE NATIONAL ENVIRONMENTAL PUBLIC HEALTH TRACKING PROGRAM.....	4
1.3 PROPOSED THREE DEMONSTRATIVE STUDIES AND OBJECTIVES	5
1.3.1 CASE-CROSSOVER ANALYSIS OF AIR POLLUTION AND CARDIO-PULMONARY HOSPITAL ADMISSIONS.....	5
1.3.2 PARTICULATE AIR POLLUTION AND TERM LOW BIRTH WEIGHT IN ALLEGHENY COUNTY	6
1.3.3 CIGARETTE SMOKING AND ADULT LEUKEMIA.....	7
2.0 CASE-CROSSOVER ANALYSIS OF AIR POLLUTION AND CARDIOPULMONARY HOSPITALIZATIONS: USING ROUTINELY COLLECTED HEALTH AND ENVIRONMENTAL DATA FOR TRACKING.....	8
2.1 ABSTRACT.....	9
2.2 INTRODUCTION	10
2.3 MATERIALS AND METHODS.....	12
2.3.1 STUDY POPULATION:.....	12
2.3.2 AIR POLLUTION AND WEATHER DATA	13
2.3.3 DATA ANALYSIS.....	14

2.4	RESULTS	15
2.5	DISCUSSION.....	17
2.6	REFERENCES	26
3.0	LOCAL VARIATION IN PARTICULATE MATTER DURING PREGNANCY AND TERM LOW BIRTH WEIGHT IN ALLEGHENY COUNTY, PA, USA	29
3.1	ABSTRACT.....	30
3.2	INTRODUCTION	31
3.3	MATERIALS AND METHODS.....	32
3.3.1	LIVE BIRTH COHORT	32
3.3.2	AIR POLLUTION DATA AND GEOGRAPHICAL DATA.....	33
3.3.3	EXPOSURE ESTIMATION.....	33
3.3.4	STATISTICAL ANALYSIS	34
3.4	RESULTS	35
3.5	DISCUSSION.....	36
3.6	REFERENCES	45
4.0	CIGARETTE SMOKING AND THE RISK OF ADULT LEUKEMIA: RESULTS FROM THE THREE MILE ISLAND COHORT STUDY	48
4.1	ABSTRACT.....	49
4.2	INTRODUCTION	50
4.3	METHODS.....	51
4.3.1	STUDY POPULATION	51
4.3.2	SMOKING INFORMATION.....	52
4.3.3	TOTAL LEUKEMIA AND AML SUBTYPE IDENTIFICATION.....	52
4.3.4	STATISTICAL ANALYSIS	53
4.4	RESULTS	54
4.5	DISCUSSION.....	56
4.6	REFERENCES	67
5.0	SUMMARY AND CONCLUSIONS	70
5.2	ASSOCIATION OF RISK OF CARDIOPULMONARY DISEASE AND AIR POLLUTION.....	70

5.3	ASSOCIATION OF RISK OF TERM LBW AND PARTICULATE AIR POLLUTION	72
5.4	ASSOCIATION OF RISK OF LEUKEMIA AND CIGARETTE SMOKING	73
5.5	STRENGTH AND LIMITATIONS OF SECONDARY DATA.....	74
5.6	OVERALL SUMMARY AND CONCLUSION	75
APPENDIX A - RISK OF ADVERSE PREGNANCY OUTCOMES AND AIR POLLUTION.....		
		77
APPENDIX B - USING SECONDARY DATA FOR ENVIRONMENTAL PUBLIC HEALTH TRACKING: SOURCES, METHODS, AND ISSUES.....		
		85
BIBLIOGRAPHY		
		124

LIST OF TABLES

Table 2-1 Distribution of specific primary diagnoses of hospital admissions among study sample in Hazelwood area, Pittsburgh, PA.....	21
Table 2-2 Percentiles and means of age, daily air pollutants and weather variables in Hazelwood, PA	21
Table 2-3 Distribution of daily air pollutants and weather factors in two-year period before and after the plant closure in Hazelwood, PA	22
Table 2-4 ORs [†] (95% CIs) of cardiopulmonary hospitalizations for PM ₁₀ before and after the closure of the plant.....	23
Table 2-5 ORs [†] (95% CIs) of cardiopulmonary hospitalizations for SO ₂ before and after the closure of plant.....	24
Table 2-6 ORs* (95% CIs) of cardiopulmonary hospitalizations per unit change in PM ₁₀ which is fitted as a continuous variable in case-crossover analyses with different control sampling approaches.....	25
Table 3-1 Weights for estimating the accumulated exposure during pregnancy.....	42
Table 3-2 Distribution of trimester-specific PM ₁₀ , Allegheny County PA, 1994-2000	42
Table 3-3 Distribution of characteristics of term singleton births by low birth weight status.....	43
Table 3-4 Risks for term low birth weight ^a according to trimester-specific exposure to PM ₁₀ ...	44
Table 4-1 Distribution of baseline characteristics among smoking groups.....	62

Table 4-2 Distribution of lymphatic and hematopoietic neoplasm incident cases occur during study period (1979-1995).....	63
Table 4-3 Distribution and risk of total leukemia among each status of cigarette smoking.....	64
Table 4-4 Distribution and risk of AML among each status of cigarette smoking.....	65
Table A-1 Results of studies of ambient air pollutants and LBW and VLBW.....	79

LIST OF FIGURES

Figure 3-1 Inverse distance interpolation.....	41
Figure 4-1 Association between risk of AML and the number years of smoking.....	66
Figure B-1 Environmental agent and adverse health outcome.....	87
Figure B-2 General procedure of data linkage.....	103

PREFACE

I owe special thanks to all of my committee members for their great support and guidance during the dissertation process. Without their personal involvement, this work could not successfully have been done so well and so quickly. I would like to thank my advisor, Dr. Evelyn Talbott, for her patience, availability and advice. Without her great support and direction, I would not have been able to accomplish this work. I am grateful for the opportunities she has provided to me. I would also like to thank everyone in my research team, especially Dr. Jeanne V. Zborowski, whose professional suggestions and editorial support did me tremendous favor. I appreciate their support more than I can say. I am very thankful to my family and friends for their emotional and editorial support to my work.

1.0 INTRODUCTION

The environment plays an essential role in the health of community. An enormous number and variety of chemicals have been introduced into the environment with the process of industrialization and economic development. They have caused considerable concerns among the community. Exposures to environmental hazards including chemical, biological and physical agents have been linked to various adverse health outcomes such as cancer (Jurewicz and Hanke 2006; Navarro Silvera and Rohan 2007), cardiovascular diseases (Delfino et al. 2005; Monarca et al. 2006), respiratory diseases (Kunzli and Tager 2005), adverse pregnancy outcomes (Perera et al. 2005; Shi and Chia 2001; Sram et al. 2005) and others in the research literature. Efficiently tracking potential health effects associated with environmental exposure appears important, urgent and necessary.

1.1 ENVIRONMENTAL EXPOSURES AND CHRONIC DISEASES

Most chronic diseases are caused by the interaction between gene and environment, i.e. diseases result from an unfavorable combination of genetics and environment (Olden and Guthrie 2001). Population-based twin studies, which can distinguish between the contributions of genetics and environment, suggest that the environment plays a prominent role in disease development (Mucci et al. 2005; O'Brien 2000). A recent study by Lichtenstein et al found that

genetics only account for 21-42% of the risk for developing the 10 common cancers studied and shared environmental exposures could account for 58-82% of the risk (Lichtenstein et al. 2000).

Environmental hazards are found not only in chemical agents used in agriculture and industrial emissions but also in household commercial products, food additives, medicine, research lab as well as natural events such as forest fire and volcano eruption. Scientific research has provided a better understanding of some environmental exposures that could cause a variety of adverse health outcomes such as arsenic and cancers in different systems (Ferreccio and Sancha 2006; Tapio and Grosche 2006), radon and lung cancer (Alavanja et al. 1994; Neuberger and Gesell 2002; Pershagen et al. 1992), lead and cognitive function (Banks et al. 1997; Rice 1996), and particulate matter and cardiopulmonary diseases (Peters and Pope 2002; Pope 2000; Pope et al. 2004). However, for many other environmental hazards, scientific evidence is less conclusive.

Billions of pounds of toxic agents are released into the environment every year where we live, eat, drink, work and play. In the report of National Survey of Public Perceptions of Environmental Health Risks, 90% of Americans believed that environmental factors like pollution, waste and chemicals are important contributors to diseases (<http://healthyamericans.org/reports/files/survey0620.pdf>). To address the public concerns, it is very important to systematically and efficiently identify environmental hazards that cause adverse health conditions such as mortality or morbidity in order to remove or minimize the health impacts of the toxic agents in the environment.

1.2 NATIONAL ENVIRONMENTAL PUBLIC HEALTH TRACKING NETWORK

Health tracking systems can be classified into two categories: passive and active tracking system. Passive tracking systems are usually based on a registry where local health care providers routinely report every new-onset case of a disease. Active tracking system mostly relies on population-based survey methods which include conducting personal interviews and reviewing medical records to collect information about personal characteristics, health conditions and others. Tracking in Environmental health is a strategy for the identification of hazards of the environment that cause death, disease, or disability, in order to facilitate the goal of prompt removal or reduction of exposures to the offending agents (Hertz-Picciotto 1996).

1.2.1 THE CURRENT STATUS OF ENVIRONMENTAL HEALTH TRACKING SYSTEMS

The current tracking systems in environmental health are specific only on either environmental hazards or health outcomes monitoring. For example, federal and state environmental agencies have routinely collected much environmental hazard data on air pollution, water pollution and others. In addition, federal, state and local public health agencies have developed registries focusing on health outcomes like cancer, birth defects as well as asthma etc. Some familiar examples of health effects datasets include the Surveillance, Epidemiology, and End Results (SEER) program, the Behavioral Risk Factor Surveillance System (BRFSS), the Iowa Birth Defects Registry, National Health Interview Survey (NHIS), and vital statistics data (See Appendix B: using secondary data for environmental public health tracking). However, few current tracking systems in environmental health can effectively

communicate with other tracking systems, efficiently track environmental exposures, and link environmental contaminants with health endpoints.

In 2000, after 18 months of review, the Pew Environmental Health Commission stated that the current system does not have the capability to respond adequately to environmental threats and that the environmental public health system was fragmented, neglected and ineffective. The commission recommended establishing a national environmental public health tracking (EPHT) network to link information on environmentally related diseases, human exposures and environmental hazards (McGeehin et al. 2004).

1.2.2 THE NATIONAL ENVIRONMENTAL PUBLIC HEALTH TRACKING PROGRAM

The Centers for Disease Control and Prevention (CDC) is leading the initiative to build a national EPHT network which is the ongoing systematic collection, integration, analysis, interpretation, and dissemination of data about environmental hazards, exposure to environmental hazards, and health effects potentially related to exposure to environmental hazards (McGeehin et al. 2004). The establishment of this network is a vital step to address the public health needs of the United States, track chronic diseases as well as environmental exposures and eventually reduce the burden of disease on the nation's population.

Currently there are 17 states and local health departments and four schools of public health funded to conduct activities that will form the basis of the tracking network. As one of four academic partners funded by CDC, the University of Pittsburgh/Drexel Academic Consortium for Excellence in Public Health Tracking (UPACE-EPHT) is uniquely positioned to interact with health-related agencies in a local, state and regional initiative to facilitate environmental capacity

building, to evaluate existing surveillance methodologies and to develop innovative strategies and tools to link hazards, exposures and health effects databases and conduct demonstrative studies of using secondary data to assess the potential associations between environmental hazards and health outcomes. The purpose of the Academic Partners program is to provide expertise and support to the National Environmental Public Health Tracking Program (NEPHTP) in the development and utilization of the Tracking Network.

1.3 PROPOSED THREE DEMONSTRATIVE STUDIES AND OBJECTIVES

One of the important activities of the academic partners in the CDC EPHT program is to conduct pilot/demonstration epidemiological studies using existing hazard, exposure and health outcome data. The main aims of this work are to examine the potential associations between environmental hazards/exposures and health outcomes and to explore methods of data linkage and data analyses. The techniques employed in the demonstration studies will be applicable to state and/or local environmental health tracking agencies. In this project, three pilot studies have been developed using secondary data from state or local environmental public health tracking agencies

1.3.1 CASE-CROSSOVER ANALYSIS OF AIR POLLUTION AND CARDIO-PULMONARY HOSPITAL ADMISSIONS

A case-crossover analysis is performed to assess the association between air pollution and cardiopulmonary hospital admissions in the Hazelwood neighborhood of Pittsburgh, PA. The

case-crossover design was first proposed by Maclure in 1991(Maclure 1991). The design is an attractive approach to estimating the effects of triggers of acute health outcomes by environmental exposure. In the design, only cases are involved and the exposure of each case during an at-risk “hazard period” just before the event is compared with the level on one or more reference days when the event did not occur. This self-matching approach will control for all known or unknown time-invariant confounders by design. This method is an alternative to time-series analysis in air pollution studies. The secondary datasets used in the study included the hospital admission data from the Pennsylvania Health Care Cost Containment Council, the climatic data of Pittsburgh from the U.S. National Climatic Data Center database and the air pollution data (PM₁₀ and SO₂) in the Hazelwood monitor site from the Allegheny County Department of Health. The objectives of this study are to assess the health effects of air pollution on cardiopulmonary disease hospitalizations, to evaluate the effects of elimination of a major point source of air pollution, i.e. the closure of LTV coke plant, on these relationships and to examine how the different methods of control sampling in THE case-crossover design influence the results (See SECTION 2).

1.3.2 PARTICULATE AIR POLLUTION AND TERM LOW BIRTH WEIGHT IN ALLEGHENY COUNTY

Low birth weight has been reported to influence the subsequent health status of individuals including morbidity in adulthood (Barker 1995; Hales 1997). Studies across the world have consistently provided evidence of the association between particulate matter and adult health including mortality and morbidity (Chen et al. 2007; Dockery et al. 1993; Kan et al. 2007; Kettunen et al. 2007; Pope 2000; Saldiva et al. 1995). A special concern of adverse health

effects of particulate air pollution on fetal health is generated. This study focuses on the associations between exposures to particulate air pollution (PM₁₀) and term low birth weight infant and further examines the effect period of particulate matter during pregnancy. The birth registration data from the Allegheny County Health Department and the air pollution data from RAND's Center for Population Health and Health Disparities (CPHHD) are obtained in this study. (See SECTION 3)

1.3.3 CIGARETTE SMOKING AND ADULT LEUKEMIA

Benzene, a well-established carcinogen and leukemogen, has been shown to be present in cigarette smoke. Wallace observed that approximately 90% of personal exposure to benzene in the United States is due to smoking (Wallace 1996). The study was designed to assess the association between cigarette smoking and leukemia, as well as its subtype, acute myeloid leukemia (AML) and to further examine the dose-response relationship between health endpoints and smoking with the number of cigarettes per day, the number of years of smoking and pack years. The TMI cohort, assembled by the Pennsylvania Department of Health to evaluate the adverse health effects of exposure of low-level radiation emitted from the TMI nuclear power plant accident on 28 March 1979, and the PA cancer registry data are obtained in this study.

The TMI cohort and PA cancer registry data was linked by the key variable of social security number (SSN). The records from these two databases refer to the same individual if the SSN is identical. This is a simple example of deterministic data linkage strategy. Due to confidentiality, this part of the work was done by the PA Department of Health. Survival analysis is performed to estimate the associations after controlling for other important confounding factors.

**2.0 CASE-CROSSOVER ANALYSIS OF AIR POLLUTION AND
CARDIOPULMONARY HOSPITALIZATIONS: USING ROUTINELY COLLECTED
HEALTH AND ENVIRONMENTAL DATA FOR TRACKING**

Xiaohui Xu, Ph.D.¹

Evelyn O. Talbott, Dr.P.H., M.P.H.¹

Jeanne V. Zborowski, Ph.D.¹

Vincent C. Arena, Ph.D.²

Judy Rager, M.P.H.¹

1. Department of Epidemiology, University of Pittsburgh Graduate School of Public Health,
Pittsburgh, PA 15261

2. Department of Biostatistics, University of Pittsburgh Graduate School of Public Health,
Pittsburgh, PA 15261

Manuscript in preparation

2.1 ABSTRACT

Background: From the early 1900s until its closure in February 1998, the LTV Steel coke oven in Pittsburgh (Hazelwood), Pennsylvania was a key source of air pollution. A case-crossover study was performed to assess the associations between daily air pollution and cardiorespiratory (ICD-9: 390-519) hospitalizations before and after plant closure and to evaluate how closure influenced these associations.

Methods: Air pollution data, climatic data and cardiorespiratory hospitalizations among Hazelwood-area residents 65 years and older were obtained for the period of 1996 through 2000. Data were analyzed using a case-crossover design and conditional logistical regression. Two distinct referent sampling approaches were compared.

Results: Significant associations were observed between the fourth quartile in PM_{10} and cardiorespiratory hospitalizations (OR: 1.12; 95%CI: 1.02-1.23) as well as cardiovascular hospitalizations only (ICD-9: 390-459) (OR: 1.13; 95%CI: 1.01-1.26) before the plant closure. After closure of the plant, PM_{10} was not significantly associated with cardiorespiratory or cardiovascular disease hospitalizations. Moreover, the referent sampling approaches did not greatly alter the estimations in the case-crossover analysis.

Conclusion: Existing secondary data is an economical source to assess the impact of point source polluters on the environmental landscape. The findings suggest that closure of the LTV coke plant was associated with a reduction risk of the cardiorespiratory and cardiovascular hospitalizations.

Key word: Case-Crossover Analysis Cardiopulmonary Air Pollution

2.2 INTRODUCTION

The environment plays an important role in health and human development. There are few surveillance systems that have the capability to adequately assess environmental threats at the state or national levels in the United States. To address this gap, in 2002 the Centers for Disease Control and Prevention (CDC) began to establish the National Environmental Public Health Tracking (EPHT) Network, which is the ongoing collection, integration, analysis, interpretation, and dissemination of data on environmental hazards; exposure to those hazards; and related health effects (1). A hallmark of the Tracking Network is the use of existing data in evaluating the relationships between exposures and health outcomes. In the present study, we demonstrated the use of data for purposes other than for which they were collected such as administrative purposes, applied available methods for linking and analyzing environmental and health outcome data, and made a positive contribution to closing gaps in methodology for conducting similar investigations within the context of the environmental public health tracking program.

Epidemiological studies consistently show that air pollutants are linked to cardiovascular and respiratory diseases (2-5). There are several potential biological mechanisms by which air pollutants cause adverse health effects. One of the potential mechanisms is related to inflammation and oxidative stress. High concentrations of oxidants and pro-oxidants contained in ambient air pollution, such as transition metals, reactive organic compounds and gases such as ozone (O₃) or nitrogen oxides (NO, NO₂) can generate reactive oxygen species(ROS)(6), alter function of mitochondria(7) or NADPH-oxidase (8), activate inflammatory cells capable of generation of ROS and reactive nitrogen species(9), and promote oxidative stress and respiratory inflammatory responses, which lead to DNA damage(10). Studies show that air pollution can increase the level of blood viscosity and modify the adhesive properties of red blood cells(11),

increase blood pressure(12) and C-reactive protein level(13), and increase heart rate(14), which lead to an increased risk of cardiovascular diseases. In addition, air pollutants, interacting with lung receptors, can activate pulmonary neural reflexes and change the autonomic function resulting in an decreased heart rate variability (HRV)(15) which partly explains the observed cardiovascular effects.

Coke plants produce a by-product of coal that burns more intensely and is used in the making of steel. These plants, as a major source of air pollution, emit gaseous air pollutants such as sulfur dioxide (SO₂), nitrogen oxide (NO₂), carbon monoxide (CO), airborne particulate matter which contains organic compounds like benzo[*a*]pyrene (B[*a*]P) and other polycyclic aromatic hydrocarbons (PAHs) (16, 17) and inorganic compounds like lead and mercury as well as other hazardous pollutants. These emissions have been linked to cardiorespiratory diseases and threaten the natural environment and overall health of the human population(18, 19).

Hazelwood is a neighborhood located in Pittsburgh, Pennsylvania. The LTV Corporation coke plant, a major local source of air pollution, was operational in the area for more than half a century. Toxic emissions from the plant led to the deterioration of the surrounding environmental landscape in the Hazelwood, Greenfield, and Squirrel Hill neighborhoods. On February 28, 1998, after eighty years of operation, the LTV plant was closed. The plant's auxiliary facilities and much of the plant itself were demolished in early 1999. The closure of the plant provided a unique opportunity to conduct a “retrospective natural intervention” study to examine whether reductions of ambient pollution concentration related to plant emissions reduced the adverse cardiorespiratory health effects in the area.

The case-crossover design was first proposed by Maclure in 1991(20). The design is an attractive approach to estimating the effects of triggers of acute health outcomes by

environmental exposure and has subsequently been widely used in air pollution studies(21-25). In this design, only cases of events are involved. The exposure of each case during an at-risk “hazard period” just before the time of the event is compared with exposure during one or more periods when the event did not occur (control periods). This self-matching approach controls by design for all known or unknown time-invariant confounders.

In the present study, a case-crossover analysis was performed to assess the association between air pollution and cardiorespiratory hospital admissions before and after the closure of this plant among the population 65 years of age and older who lived in the Hazelwood area, and to evaluate how the closure of a major point source of air pollution influences the associations. We also compared two different strategies of selecting referent periods in the case crossover analysis and evaluated how the selection of referent periods influenced the results.

Existing data used in the study included hospital admission data from the Pennsylvania Health Care Cost Containment Council, climatic data of Pittsburgh from the U.S. National Climatic Data Center Database, and a pollution monitoring data from the Allegheny County Health Department.

2.3 MATERIALS AND METHODS

2.3.1 STUDY POPULATION:

The Pennsylvania Health Care Cost Containment Council (PHC4), Special Request Unit (Harrisburg, PA) provided the cardiorespiratory hospital admission data for subjects 65 years of age and older in Hazelwood and surrounding neighborhoods in the city of Pittsburgh,

Pennsylvania. The study location included all zip codes surrounding the LTV coke plant with the plant as the center, i.e. zip codes 15207, 15120, 15217 and 15218. A dataset that included all inpatient cardiovascular and respiratory (i.e. Cardiorespiratory) hospital admissions in the Hazelwood area between 1996 and 2000 was created. The records contained a primary discharge diagnosis of the circulatory system [International Classification of Disease, 9th Revision (ICD-9) codes of 390-459] or respiratory system (ICD-9 codes of 460-519]. The study was limited to the first event for those who were hospitalized for Cardiorespiratory diseases in a one month period.

2.3.2 AIR POLLUTION AND WEATHER DATA

Ambient air levels of specific criteria pollutants were obtained in electronic format for the Hazelwood monitoring site from December 1995 through January 2001. The data collected included 24-hour values for sulfur dioxide (SO₂) and PM₁₀. The daily minimum, maximum and mean values of SO₂ and PM₁₀ were computed for the present study. These daily measurements were assumed to represent the exposure level for the entire Hazelwood study area.

Daily meteorological data were obtained from the U.S. National Climatic Data Center database from the monitoring site at the Pittsburgh International Airport, Allegheny County (Coopid: 366993, Wbandid: 94,823, Latitude: 40°30', Longitude:-80°14'). The information for daily mean temperature and daily mean relative humidity was abstracted from the database for the time period of 1996-2000.

2.3.3 DATA ANALYSIS

Data were analyzed with the case-crossover technique which is an alternative to time-series analysis for assessing acute health effects of air pollution. In the design, cases serve as their own controls. A subject's exposure at the time of a health event (case-period) is compared with exposures at previous or subsequent points of time when that subject was a non-case (control-period). In the present study, the case-period is defined as the date of hospitalization for any cardiorespiratory diseases. The control periods were chosen using two methods, either the bidirectional control sampling approach or the time-stratified sampling approach in order to control relevant time-varying confounders (26). In the bidirectional control sampling approach, the control periods were selected as 7 and 14 days before and after the date of hospitalization in order to control for day-of-week effect. In the time-stratified method, the stratum is defined as the month of event. All other same days of week as the case period in the month were selected as the control periods. For example, if a case occurs in March on a Wednesday, all other Wednesdays in March are the control periods.

In the bidirectional control sampling approach, the associations between air pollutants and hospitalization for cardiovascular diseases only or respiratory diseases only as well as both combined (cardiorespiratory) were estimated for the time period January 1, 1996 to February 28, 1998 (before closure of the plant) and for the period March 1, 1998 to December 31, 2000 (after closure of the plant), respectively. The SAS conditional logistic regression procedure (Proc Phreg) was applied to estimate the associations between hospital admissions and air pollution. Air pollutants were fitted into the model as continuous variables or categorical variables (quartiles). Odds ratios (ORs) were calculated for continuous variables and the quartiles of each

pollutant during each time period with the first quartile as a reference after adjusting for daily meteorological factors including current day's mean temperature and relative humidity.

In the time-stratified control sampling approach, all analyses were performed with the Case-Crossover Analysis Tool (Beta V1.1 C-CAT), developed by Apex Epidemiology Research, LLC, in conjunction with the New York State Department of Health, to provide an easy-to-use interface to SAS software that implements time-stratified case-crossover analysis. Air pollutants were fitted into the model only as continuous variables. The results of this approach were compared with the previous approach. In addition, the impact of different referent period selections on the results was also examined.

2.4 RESULTS

Among the residents of Hazelwood area who were at least 65 years of age, there were a total of 13,437 cardiorespiratory disease admissions (9,935 cardiovascular disease admissions and 3,502 respiratory diseases admissions) during 1996-2000. Table 1 summarizes the number of admissions that had a specific diagnosis of cardiovascular disease and respiratory disease before and after the closure of the LTV coke plant. There were 4,491 and 5,444 cardiovascular disease hospitalizations before and after the closure of the plant, respectively. In addition, 1,596 and 1,906 respiratory disease hospitalizations were respectively observed before and after the closure of the plant in this study.

The distributions of air pollutants, age and meteorological variables before and after the closure of plant are described in Table 2-2. The distribution of age of hospitalizations is similar before and after the plant closure. The percentiles (10th, 25th, 50th, 75th and 90th) of PM₁₀ and

SO₂ were all higher before than after the closure of the plant. The daily averages of PM₁₀ were 27.75 µg/m³ before and 21.96 µg/m³ after the plant closed, respectively. The average of SO₂ decreased to nearly half of the previous levels (from 16.4 ppb to 8.7 ppb) after the plant was closed. For considering the difference of length of time between before and after the plant closure, air pollutants and climatic factors are compared for two-year period before and after the plant closure, respectively. The results are described in Table 2-3. The change of air pollutants is the same regardless of the difference of length of time. The percentiles of daily temperature and daily relative humidity were similar in two-year period before and after the closure of the plant.

Table 2-4 shows the associations between quartiles of PM₁₀ and cardiorespiratory hospitalizations after controlling for temperature and relative humidity. The results showed significant associations between the fourth quartile in PM₁₀ and cardiorespiratory hospitalizations (OR: 1.12; 95%CI: 1.02-1.23) as well as cardiovascular disease hospitalization only (OR: 1.13; 95%CI: 1.01-1.26) before the closure of the plant. Moreover, significant trends for increasing risks on cardiorespiratory hospitalizations (p=0.008) and cardiovascular disease hospitalizations (p=0.02) with increasing concentration of PM₁₀ were observed. After the closure of the plant, the associations with cardiorespiratory hospitalization were no longer statistically significant. In this study, no significant association between respiratory disease hospitalization and PM₁₀ was found either before or after the plant closure.

The associations between SO₂ and cardiorespiratory hospitalizations were illustrated in Table 2-5. There were no significant associations between SO₂ and hospitalizations of cardiovascular diseases only or respiratory diseases only or both combined either before or after the closure of the plant.

Table 2-6 shows the associations between PM₁₀ and cardiorespiratory hospitalizations with different referent periods in the case-crossover analysis. The results from the bidirectional control sampling approach are similar to the results of the time-stratified control sampling method.

2.5 DISCUSSION

Few studies are currently available to assess how the addition or elimination of a single point major air pollution source affects the risk of adverse health effects in specific geographic area. This study, for the first time, used a case-crossover design to evaluate this relationship using existing secondary data. The case-crossover approach is a self-matched case control study, which has great advantages in comparison with a time series approach. This approach controls certain confounding factors by design rather than by modeling. With this design, personal characteristics such sex, age, race and time-invariant variables are controlled by the design. By using symmetric 7 and 14 day reference periods, the case-crossover design eliminates the confounding effects of long-term trends, seasonality and day of week(27). Therefore, the use of a case-crossover design avoids common concerns about the complex mathematical modeling and adequacy of seasonal control.

This study demonstrated that the levels of PM₁₀ and SO₂ decreased in Hazelwood area after the plant was closed. PM₁₀ declined an average of 6 µg/m³ and SO₂ was reduced to half of the level when the plant was open. Significant associations between PM₁₀ and cardiovascular diseases as well as cardiorespiratory disease hospitalizations were observed before the closure of plant. Moreover, a trend of increased risk for cardiovascular disease as well as cardiorespiratory

disease hospitalizations with increasing particulate matter was observed before the plant closure in the study. Interestingly, after the closure, the associations were no longer statistically significant. Sulfur dioxide did not show a significant association with any outcomes in this study either before or after the closure of the plant. Because sulfur dioxide, a highly soluble gas, can be easily and predominantly stripped out of the upper airways(28), the role of sulfur dioxide in developing adverse health effects is less coherent. Several case-crossover studies reported that there are no significant associations between sulfur dioxide and respiratory disease(22, 29) as well as cardiovascular disease hospitalization(30, 31).

The reduced health effects of PM₁₀ after the plant closure could be due to two facts: the change in composition and/or concentration of particulate matter. After the elimination of the LTV coke plant, some toxic elements in the particulate matter associated with the plant emission have also been removed. The monitored data from the Allegheny County Health Department showed that the concentrations of benzo(α)pyrene [B(α)P], as a recognized carcinogen as well as respiratory, endocrine and immunological toxicant, was significantly lower after the plant closure. B(α)P can not however be used as a predictor in the study because it was measured every 7 days. Another potential fact is the reduced concentration of particulate matter in the area. At the low level of particulate matter, there might be no adverse health effect on cardiorespiratory hospitalizations or the association is too weak to be detected on this population sample size.

The findings of this study support that particulate air pollution could increase the risks of cardiorespiratory hospitalizations before the LTV coke plant closure and an elimination of a major air pollution source could lead to a reduced risk of adverse health effects, which is consistent with other epidemiological studies. Pope et al conducted a cross-sectional study to

assess how the operation of a steel mill influences the association between hospital admissions and PM₁₀ in Utah Valley during the period April 1985-February 1988. The results showed that children's hospital admissions were two to three times higher during the winters when the mill was open compared to when it was closed (32). The findings in this study further strengthen the epidemiologic evidence of acute adverse health effects of airborne particle levels.

In this study, we also evaluated how the control sampling approaches influence the estimations in the case-crossover analysis. PM₁₀ was the only pollutant used to evaluate these approaches because sulfur dioxide had no associations with the outcomes in this study and Akaike's information criterion (AIC) was increased in both methods after SO₂ was fitted as a continuous variable, suggesting a poor model fit. The estimations from the bidirectional control sampling methods are similar to those from the time-stratified method. The results of the comparison suggested that the control sampling methods in case-crossover analysis did not alter the conclusion.

A potential limitation of this study is that only one air pollution monitoring site was available in the Hazelwood area. Individual exposure was estimated by using the measurement from this fixed outdoor monitoring station. The potential misclassification of exposure due to lack of personal measurement can not be avoided. However, the bias might underestimate the association and shift the results toward the null (33, 34). In addition, we did not assess the promoting role of other pollutants such as nitrogen dioxide or ozone because the data of these pollutants from the monitor site were not available or complete. However, particulate matter is likely a reasonable indicator of other pollutants due to the high collinearity between particulate matter and other pollutants (35). Another limitation is potential for selection bias because we were unable to include all cardiorespiratory disease cases in the study if unreported to the PHC4.

Finally, another limitation of this study is that the sample size of respiratory disease hospitalization might be not large enough to have sufficient power to detect a statistically significant difference.

In summary, this study demonstrates the utility of secondary datasets and the methodology of the case-crossover approach to evaluate the associations between environmental air pollutants and acute health outcomes. This investigation demonstrated important strategies for using existing data to address different health issues and developing the methodology of data linkage and data analysis in conducting similar studies in an environmental public tracking program. Moreover, this analysis provides, for the first time, evidence of a reduction in air pollution resulting in a decrease in risk of Cardiorespiratory disease hospitalization using a case-crossover design. The findings reinforce the deleterious impact of air pollution and provide new epidemiologic information that the elimination or addition of a major point source of pollution could change a local environmental landscape and influence the health of the population living in the area. Further studies are required to confirm these findings.

Table 2-1 Distribution of specific primary diagnoses of hospital admissions among study sample in Hazelwood area, Pittsburgh, PA

Diagnosis	LTV Plant open (1/1/1996-02/28/1998)	LTV Plant closed (3/1/1998-12/31/2000)	Total
Cardiovascular	4,491	5,444	9,935
Respiratory	1,596	1,906	3,502
Combined	6,087	7,350	13,437

Table 2-2 Percentiles and means of age, daily air pollutants and weather variables in Hazelwood, PA

Variables	Time periods	10%	25%	50%	75%	90%	Mean
Age (years)	LTV Plant open	68	72	77	83	88	77.4
	LTV Plant closed	68	72	78	83	88	78.0
PM ₁₀ , µg/m ³	LTV Plant open	13.1	17.8	24.0	35.0	46.8	27.8
	LTV Plant closed	10.4	13.5	19.1	27.7	37.2	22.0
SO ₂ , ppb	LTV Plant open	4.7	8.2	14.1	22.4	30.1	16.4
	LTV Plant closed	3.4	5.0	7.7	11.0	15.3	8.7
Temperature (°F)	LTV Plant open	26.0	35.0	48.0	65.0	72.0	48.8
	LTV Plant closed	28.0	40.0	55.0	67.0	74.0	52.5
Humidity (%)	LTV Plant open	30.0	40.0	49.0	62.0	73.0	51.1
	LTV Plant closed	29.0	37.0	48.0	60.0	72.0	49.1

Table 2-3 Distribution of daily air pollutants and weather factors in two-year period before and after the plant closure in Hazelwood, PA

Variables	Time periods	Mean	10%	25%	50%	75%	90%
PM ₁₀ , µg/m ³	LTV Plant open	28.2	13.5	18.2	24.6	35.4	47.4
	LTV Plant closed	21.3	10.2	13.1	18.5	26.9	36.3
SO ₂ , ppb	LTV Plant open	16.1	4.5	8.1	13.8	22.2	29.5
	LTV Plant closed	8.5	3.1	4.9	7.7	10.9	14.5
Temperature (°F)	LTV Plant open	50.4	27	37	51	66	73
	LTV Plant closed	51.6	26	38	55	67	73
Humidity (%)	LTV Plant open	50.5	30	40	49	61	72
	LTV Plant closed	48.5	28	36	48	60	71

Table 2-4 ORs[†] (95% CIs) of cardiopulmonary hospitalizations for PM₁₀ before and after the closure of the plant

Hospitalizations	Exposure levels	PM ₁₀	
		LTV Plant open (01/01/1996- 02/28/1998)	LTV Plant closed (03/01/1998-12/31/2000)
Cardiovascular and Respiratory diseases combined	<25 th percentile *	1.0	1.0
	25 th to 50 th percentile	0.99 (0.91-1.08)	1.03 (0.96-1.11)
	50 th to 75 th percentile	1.08 (0.99-1.18)	1.02 (0.94-1.10)
	≥75 th percentile	1.12 (1.02-1.23) §	0.96 (0.88-1.05)
	P for trend	0.008 [§]	0.61
Cardiovascular diseases only	<25 th percentile *	1.0	1.0
	25 th to 50 th percentile	1.01 (0.91-1.11)	1.09 (0.99-1.81)
	50 th to 75 th percentile	1.10 (0.99-1.21)	1.04 (0.95-1.14)
	≥75 th percentile	1.13 (1.01-1.26) §	0.94 (0.85-1.04)
	P for trend	0.02 [§]	0.42
Respiratory diseases only	<25 th percentile *	1.0	1.0
	25 th to 50 th percentile	0.97 (0.82-1.39)	0.89 (0.78-1.03)
	50 th to 75 th percentile	1.05 (0.89-1.24)	0.97 (0.84-1.13)
	≥75 th percentile	1.11 (0.92-1.35)	1.03 (0.87-1.22)
	P for trend	0.22	0.71

* Reference group

† Odds ratios adjusted for daily weather variables (daily mean temperature and humidity)

§ P<0.05

Table 2-5 ORs[†] (95% CIs) of cardiopulmonary hospitalizations for SO₂ before and after the closure of plant

Hospitalizations	Exposure levels	SO ₂	
		LTV Plant open (01/01/1996- 02/28/1998)	LTV Plant closed (02/28/1998- 12/31/2000)
Cardiovascular and Respiratory diseases combined			
	<25 th percentile *	1.0	1.0
	25 th to 50 th percentile	1.00 (0.92-1.09)	1.03 (0.95-1.11)
	50 th to 75 th percentile	1.06 (0.97-1.16)	0.99 (0.92-1.07)
	≥75 th percentile	1.01 (0.93-1.11)	1.02 (0.94-1.11)
	P for trend	0.47	0.68
Cardiovascular diseases only			
	<25 th percentile *	1.0	1.0
	25 th to 50 th percentile	1.04 (0.95-1.15)	1.03 (0.96-1.12)
	50 th to 75 th percentile	1.07 (0.96-1.19)	0.96 (0.88-1.05)
	≥75 th percentile	1.05 (0.94-1.17)	0.99 (0.89-1.09)
	P for trend	0.25	0.59
Respiratory diseases only			
	<25 th percentile *	1.0	1.0
	25 th to 50 th percentile	0.89 (0.75-1.05)	1.02 (0.88-1.18)
	50 th to 75 th percentile	1.05 (0.88-1.24)	1.09 (0.94-1.27)
	≥75 th percentile	0.92 (0.77-1.10)	1.12 (0.95-1.32)
	P for trend	0.59	0.09

* Reference group

† Odds ratios adjusted for daily weather variables (daily mean temperature and humidity)

Table 2-6 ORs* (95% CIs) of cardiopulmonary hospitalizations per unit change in PM₁₀ which is fitted as a continuous variable in case-crossover analyses with different control sampling approaches

Hospitalizations	Time periods	Control sampling approach	
		Bidirectional control sampling	Time-stratified control sampling
Cardiovascular and Respiratory diseases combined	LTV Plant open	1.003 (1.001-1.005)	1.002 (1.00-1.005)
	LTV Plant closed	0.998(0.995-1.001)	0.997 (0.994-1.00)
Cardiovascular diseases only	LTV Plant open	1.003 (1.00-1.006)	1.003(1.00-1.005)
	LTV Plant closed	0.997 (0.994-1.00)	0.996 (0.993-1.00)
Respiratory diseases only	LTV Plant open	1.003 (0.998-1.008)	1.001 (0.996-1.006)
	LTV Plant closed	1.00 (0.995-1.006)	0.998 (0.992-1.004)

* Odds ratios adjusted for daily weather variables (daily mean temperature and humidity)

2.6 REFERENCES

1. McGeehin MA, Qualters JR, Niskar AS. National environmental public health tracking program: bridging the information gap. *Environ Health Perspect* 2004;112:1409-13.
2. Pope CA, 3rd, Thun MJ, Namboodiri MM, et al. Particulate air pollution as a predictor of mortality in a prospective study of U.S. adults. *Am J Respir Crit Care Med* 1995;151:669-74.
3. Penttinen P, Timonen KL, Tiittanen P, Mirme A, Ruuskanen J, Pekkanen J. Ultrafine particles in urban air and respiratory health among adult asthmatics. *Eur Respir J* 2001;17:428-35.
4. Magari SR, Schwartz J, Williams PL, Hauser R, Smith TJ, Christiani DC. The association of particulate air metal concentrations with heart rate variability. *Environ Health Perspect* 2002;110:875-80.
5. Baldacci S, Viegi G. Respiratory effects of environmental pollution: epidemiological data. *Monaldi Arch Chest Dis* 2002;57:156-60.
6. Fubini B, Fenoglio I, Ceschino R, et al. Relationship between the state of the surface of four commercial quartz flours and their biological activity in vitro and in vivo. *Int J Hyg Environ Health* 2004;207:89-104.
7. Li N, Sioutas C, Cho A, et al. Ultrafine particulate pollutants induce oxidative stress and mitochondrial damage. *Environ Health Perspect* 2003;111:455-60.
8. Knaapen AM, Borm PJ, Albrecht C, Schins RP. Inhaled particles and lung cancer. Part A: Mechanisms. *Int J Cancer* 2004;109:799-809.
9. Schins RP, Lightbody JH, Borm PJ, Shi T, Donaldson K, Stone V. Inflammatory effects of coarse and fine particulate matter in relation to chemical and biological constituents. *Toxicol Appl Pharmacol* 2004;195:1-11.
10. Risom L, Moller P, Loft S. Oxidative stress-induced DNA damage by particulate air pollution. *Mutat Res* 2005;592:119-37.
11. Seaton A, Soutar A, Crawford V, et al. Particulate air pollution and the blood. *Thorax* 1999;54:1027-32.
12. Brook RD. You are what you breathe: evidence linking air pollution and blood pressure. *Curr Hypertens Rep* 2005;7:427-34.

13. Donaldson K, Stone V, Seaton A, MacNee W. Ambient particle inhalation and the cardiovascular system: potential mechanisms. *Environ Health Perspect* 2001;109 Suppl 4:523-7.
14. Peters A, Perz S, Doring A, Stieber J, Koenig W, Wichmann HE. Increases in heart rate during an air pollution episode. *Am J Epidemiol* 1999;150:1094-8.
15. Pope CA, 3rd, Hansen ML, Long RW, et al. Ambient particulate air pollution, heart rate variability, and blood markers of inflammation in a panel of elderly subjects. *Environ Health Perspect* 2004;112:339-45.
16. Liberti L, Notarnicola M, Primerano R, Zannetti P. Air pollution from a large steel factory: polycyclic aromatic hydrocarbon emissions from coke-oven batteries. *J Air Waste Manag Assoc* 2006;56:255-60.
17. Wang J, Zhu L, Shen X. [PAHs pollution in air of coke plant and health risk assessment]. *Huan Jing Ke Xue* 2003;24:135-8.
18. Bhopal RS, Phillimore P, Moffatt S, Foy C. Is living near a coking works harmful to health? A study of industrial air pollution. *J Epidemiol Community Health* 1994;48:237-47.
19. Dolk H, Thakrar B, Walls P, et al. Mortality among residents near cokeworks in Great Britain. *Occup Environ Med* 1999;56:34-40.
20. Maclure M. The case-crossover design: a method for studying transient effects on the risk of acute events. *Am J Epidemiol* 1991;133:144-53.
21. Lee JT, Schwartz J. Reanalysis of the effects of air pollution on daily mortality in Seoul, Korea: A case-crossover design. *Environ Health Perspect* 1999;107:633-6.
22. Sunyer J, Schwartz J, Tobias A, Macfarlane D, Garcia J, Anto JM. Patients with chronic obstructive pulmonary disease are at increased risk of death associated with urban particle air pollution: a case-crossover analysis. *Am J Epidemiol* 2000;151:50-6.
23. Neas LM, Schwartz J, Dockery D. A case-crossover analysis of air pollution and mortality in Philadelphia. *Environ Health Perspect* 1999;107:629-31.
24. Levy D, Sheppard L, Checkoway H, et al. A case-crossover analysis of particulate matter air pollution and out-of-hospital primary cardiac arrest. *Epidemiology* 2001;12:193-9.
25. Barnett AG, Williams GM, Schwartz J, et al. The effects of air pollution on hospitalizations for cardiovascular disease in elderly people in Australian and new zealand cities. *Environ Health Perspect* 2006;114:1018-23.
26. Levy D, Lumley T, Sheppard L, Kaufman J, Checkoway H. Referent selection in case-crossover analyses of acute health effects of air pollution. *Epidemiology* 2001;12:186-92.

27. Lee JT, Kim H, Schwartz J. Bidirectional case-crossover studies of air pollution: bias from skewed and incomplete waves. *Environ Health Perspect* 2000;108:1107-11.
28. Health effects of outdoor air pollution. Committee of the Environmental and Occupational Health Assembly of the American Thoracic Society. *Am J Respir Crit Care Med* 1996;153:3-50.
29. Sunyer J, Basagana X, Belmonte J, Anto JM. Effect of nitrogen dioxide and ozone on the risk of dying in patients with severe asthma. *Thorax* 2002;57:687-93.
30. Tsai SS, Goggins WB, Chiu HF, Yang CY. Evidence for an association between air pollution and daily stroke admissions in Kaohsiung, Taiwan. *Stroke* 2003;34:2612-6.
31. D'Ippoliti D, Forastiere F, Ancona C, et al. Air pollution and myocardial infarction in Rome: a case-crossover analysis. *Epidemiology* 2003;14:528-35.
32. Pope CA, 3rd. Respiratory disease associated with community air pollution and a steel mill, Utah Valley. *Am J Public Health* 1989;79:623-8.
33. Katsouyanni K, Touloumi G, Spix C, et al. Short-term effects of ambient sulphur dioxide and particulate matter on mortality in 12 European cities: results from time series data from the APHEA project. *Air Pollution and Health: a European Approach. Bmj* 1997;314:1658-63.
34. Zeger SL, Thomas D, Dominici F, et al. Exposure measurement error in time-series studies of air pollution: concepts and consequences. *Environ Health Perspect* 2000;108:419-26.
35. Gadgil AS, Jadhav RS. Street-level concentrations of suspended particulate matter (SPM), nitrogen dioxide (NO₂) and sulphur dioxide (SO₂) in Pune City. *J Environ Sci Eng* 2004;46:143-50.

**3.0 LOCAL VARIATION IN PARTICULATE MATTER DURING PREGNANCY
AND TERM LOW BIRTH WEIGHT IN ALLEGHENY COUNTY, PA, USA**

Xiaohui Xu, Ph.D.¹

Ravi K. Sharma Ph.D.²

Evelyn O. Talbott, Dr.P.H., M.P.H.¹

Jeanne V. Zborowski, Ph.D.¹

Vincent C. Arena, Ph.D.³

Devra Lee Davis Ph.D.¹

Conrad D Volz Dr. P.H.⁴

Judy Rager, M.P.H.¹

1. Department of Epidemiology, Graduate School of Public Health,
Universith of Pittsburgh

2. Department of Behavioral Community Health Sciences,
Graduate School of Public Health, Universith of Pittsburgh

3. Department of Biostatistics, Graduate School of Public Health,
Universith of Pittsburgh

4. Department of Environmental and Occupational Health,
Graduate School of Public Health, Universith of Pittsburgh

Manuscript in preparation

3.1 ABSTRACT

Introduction: Low birth weight has been associated with increased risks of mortality and/or morbidity in childhood and adulthood. The aim of this study is to explore the association between particulate matter (PM₁₀) and term low birth weight (birth weight <2,500g).

Methods: Birth data from the Allegheny County Health Department, Pittsburgh PA and PM₁₀ air data generated with inverse-distance interpolation by RAND's Center for Population Health and Health Disparities, were obtained. The study population consisted of all term singleton live births (gestational age ≥ 37 weeks) born between Jan 1st, 1994 to Dec 31st, 2000. Infants with birth weight <2,500g were classified as LBW. Logistic regression was performed to estimate the association per inter-quartile range increase in PM₁₀.

Results: The results showed that the odds ratios of term LBW per inter-quartile range increase in PM₁₀ were 1.13 (95%CI: 1.02-1.25) during the first trimester and 1.10 (95%CI: 1.00-1.22) during the second trimester after adjusted for other important covariates, respectively.

Conclusion: The findings of the study support that exposure to PM₁₀ could increase risks of term LBW. Further studies are warrant to corroborate these findings.

Keywords: LBW PM10 air pollution Inverse-distance interpolation

3.2 INTRODUCTION

A growing body of evidence suggests that maternal exposure to air pollution is associated with adverse pregnancy outcomes. Studies conducted worldwide have investigated the health effects of air pollution on preterm delivery (PTD) (1, 2), low birth weight (3-5), intrauterine growth restriction (IUGR) (6, 7) and birth defects (8, 9). The air pollutants of concern in these studies include carbon monoxide (CO), nitrogen dioxide (NO₂), sulfur dioxide (SO₂), ozone (O₃), and particulate matter (PM).

Low birth weight (birth weight less than 2500 g) is comprised of the two overlapping etiologies of PTD as well as IUGR. Term LBW is caused by fetal growth retardation instead of early delivery. LBW has been widely reported to influence the health status of individuals, including increased mortality and morbidity in childhood (10, 11) and an elevated risk of hypertension, coronary heart disease, Type II diabetes in adulthood, abnormalities of lipid metabolism and blood coagulation (12-15). The public health relevance on this important health issue is evident.

Studies across the world have consistently provided evidence that exposure to particulate matter increases risks of mortality and morbidity among adults (16-23). There is a great concern about the association between particulate matter and fetal health outcomes, especially low birth weight. However, the findings of particulate matter in fetal health research are inconsistent, especially regarding the effect period and the strength of association of particulate matter. In addition, many prior studies did not account for important confounders such as maternal smoking, gestational age and weight gain. Previous studies have also had limited spatial

information on pollution sources and concentrations because the exposure mostly relied on the measurements at monitoring stations.

In the present study, we investigated the possible association between particulate matter (PM₁₀) and term LBW in Allegheny County, PA (USA). The specific purposes of this study are to use PM₁₀ data spatially and temporally being generated with inverse-distance interpolation, to estimate the levels of PM₁₀ corresponding to the first, second, and third trimesters of pregnancy, to evaluate the association between PM₁₀ and term low birth weight, and to explore the possible techniques of data management, data linkage and data analysis for environmental public health tracking to evaluate these relationships.

3.3 MATERIALS AND METHODS

3.3.1 Live birth cohort

All live birth data in Allegheny County, PA was obtained from the birth registry data maintained by the Allegheny County Health Department, PA for the period January 1, 1994 through December 31, 2000. The information in the database includes date of birth, birth weight, gestational age, parity, birth order, maternal age, education, race and geographical location of birth (neighborhood in the city of Pittsburgh and municipality outside the city) and other reproductive data. Individual data of all singleton live births to the Allegheny County residents were abstracted from the database and used in the present study.

3.3.2 Air pollution data and geographical data

The quarterly measures of PM_{10} in census tracts in Allegheny County were obtained from the RAND's Center for Population Health and Health Disparities (CPHHD) Data Core. The geographic-specific data were derived from the publicly available Environmental Protection Agency (EPA) Air Quality System (AQS) Criteria pollutant data with the inverse-distance interpolation based on the 1990 based census tracts. The following are the basic processes used to construct the dataset.

The hourly data of PM_{10} between 1994 and 2000 were obtained from the U.S EPA AQS. The data were aggregated from daily to quarterly levels for each monitor site. The quarterly PM_{10} measures from the EPA monitor sites were used to estimate the levels at the centroids of census tracts, which represent the levels of the census tracts. All distances from the census tract centroid to the site locations were computed geospatially using ArcGIS (Version 9.1, ESRI Inc. Redlands, WA, USA). The measures from the monitor sites within 100 kilometers between the centroid and their locations are used to estimate the level of that census tract. The estimation involves various inverse distance-weighting schemes that are a function of the distance between the census tract centroid and the PM monitors (see Figure 1). Therefore, the sites further away will have less influence and thus less contribution overall to the estimation.

3.3.3 Exposure estimation

The quarterly PM_{10} data were aggregated from census tract to neighborhood in the city or municipality outside the city in geographical level. The average of the quarterly PM_{10} measures of all census tracts in each neighborhood or municipality was used to estimate the quarterly level

of that neighborhood or municipality. It assumed that all births in the same neighborhood or municipality had the same exposure at a single time.

Air pollution estimates for each individual birth were assigned using the neighborhood or municipality of the maternal residence at infant's birth. For each birth, the accumulated first 3-month, 6-month and 9-month exposures was computed by weighting the quarterly PM10 measures based on the mother's residence (neighborhood or municipality), the month, quarter and year of birth conception. The weights of estimating the accumulated exposures are shown in the table 1. The date of conception was computed based on the date of birth and gestational age. For example, for an infant who was conceived in February 1994, the accumulated first 3-month exposure is two times the first quarterly PM10 measure in 1994 plus one time the second quarterly PM10 measure in 1994. The monthly averages of trimester-specific exposures were calculated based on the accumulated exposures.

3.3.4 Statistical Analysis

In this study, our analyses were focused on the effects of PM₁₀ on birth weight mediated by reduced fetal growth as opposed to early delivery. Therefore, the study population was restricted to infants who were born at term with gestational age ≥ 37 weeks. The relationship between particles and term low birth weight was evaluated using logistic regression analyses. Adverse pregnancy outcomes, i.e. term LBW, defined as dichotomous categories, represent dependent variables in the analysis. A term LBW infant is defined as a live-birth infant weighing $< 2,500\text{g}$ and gestational age ≥ 37 complete weeks.

Several known risk factors for term LBW that could potentially confound the relationship between LBW and air pollution were also included in regression models: maternal age (< 20 , 20-

29, 30-34,35-39, ≥ 40 years), maternal race (black, white, others), maternal education (<9, 9-11, 12, 13-15, ≥ 16 years), maternal tobacco use (yes/no), level of prenatal care (none, during first trimester, after first trimester), history of LBW or preterm infant (one or more vs. none), parity (first birth vs. second or subsequent birth), birth season, infant sex, gestational age (measured in weeks) and maternal weight gained.

Air pollution exposures were fitted into the logistic regression model as continuous variable. The odds ratios with per inter-quartile range increase in PM₁₀, i.e. and $7 \mu\text{g}/\text{m}^3$ increase, were estimated for each trimester.

3.4 RESULTS

A total of 100,595 singleton birth records were available for Allegheny County between 1994 and 2000. Among 92,447 singleton term births (≥ 37 completed weeks gestation), 47,221 were male births and 45,226 female births. The sex ration at birth is 104 per 100 females. A total of 2,058 (2.2%) low weight births occurred over the entire study period.

The mean trimester-specific exposures and 9-month period for PM₁₀ during the study period were 28.7, 28.2, 28.1 and $28.1 \mu\text{g}/\text{m}^3$ respectively. The inter-quartile range for each trimester is nearly $7 \mu\text{g}/\text{m}^3$ (table 3-2).

Table 3-3 summarizes the distributions of characteristics of singleton births among term LBW. The results of crude estimations suggested that term LBW was significantly associated with maternal age, race, education, tobacco use, infant gender, previous LBW or PTD, parity, prenatal care, gestational age and weight gained. For term LBW, younger (<20) and older (≥ 40) maternal age, no previous birth, low levels of education and prenatal care, tobacco use and

previous LBW or PTD could increase the risks of term LBW. African-American women had a higher risk of term LBW than whites. Male infant sex was negatively associated with term LBW. Risk of term LBW decreased as the increase of gestational age and maternal weight gained (See Table 3-3).

Table 4-4 illustrates the risks of term LBW associated with per inter-quartile range increase in exposure to trimester-specific PM₁₀. Significant associations between term LBW and the first trimester exposure to PM₁₀ was observed in the study after controlling for other covariates including for maternal age, maternal race, maternal education, smoking, weight gain, gender of infant, gestation age, parity, previous LBW or preterm birth, level of prenatal care and birth season. The odds ratios of term LBW for per inter-quartile range increase in PM₁₀ were 1.13 (95%CI: 1.02-1.25) during the first trimester and 1.10 (95%CI: 1.00-1.22) during the second trimester, respectively. There is no significant association between term LBW and PM₁₀ exposure during the third trimester (OR=1.05, 95%CI: 0.96-1.16) and 9-month period (OR=1.07, 95%CI: 0.99-1.14).

3.5 DISCUSSION

The sex ratio at birth is one of the stable parameters of a population and it should lie in a narrow range of 100-108 males per 100 females. In this study, the sex ratio is 104 per 100 females, which is similar to the national level in the USA (24). We examined associations between term low birth weight and exposure to PM₁₀ at various stages of pregnancy. Increased risks of LBW were observed for mother's exposure to PM₁₀ during the first trimester after adjusting for other important confounding factors including maternal age, race, education,

smoking, weight gain, infant gender, gestational age, parity, history of LBW or PTD, prenatal care and birth season. The findings of this study support that PM₁₀ exposure could increase risk of term LBW.

In adverse pregnancy outcome research, it is important to identify the critical time points during pregnancy when exposure to air pollutants might be most harmful (25). Our finding of an adverse effect of ambient particulates in the first trimester on pregnancy outcomes is consistent with some previous studies. Dejmek et al conducted a study in the Treplice District of the Czech Republic and revealed that exposure to PM₁₀ during the first month was associated with intrauterine growth retardation (IUGR) (6). The study from Seoul found that carbon monoxide, nitrogen dioxide, sulfur dioxide, and total suspended particle concentrations in the first trimester of pregnancy period are risk factors for low birth weight (26). The finding from a study in Brazil also showed that exposure from the first trimester was the most important for LBW(27). However, other studies found that exposures during the third trimester were importantly linked with birth weight. Chen et al found that exposure to PM₁₀ in the third trimester of pregnancy was negatively associated with birth weight in North Nevada, USA(28). Wang and colleagues reported that concentrations of TSP and SO₂ in the last trimester of pregnancy were associated with low birth weight in Beijing, China(29). A study in Poland also found that PM_{2.5} exposure in the second trimester of pregnancy was negatively associated with birth weight (30). Therefore exposure during the earlier or later stage of pregnancy could present the possibility of interference with final infant weight gain.

Although the specific biological mechanisms that may account for the association between ambient air pollution and adverse pregnancy outcomes are not well known, the effect of particulate matter exposure during pregnancy on adverse pregnancy outcomes has a plausible

biological basis. It has been assumed that prenatal exposure to ambient air pollution might be similar to maternal smoking which can result in some adverse reproductive outcomes. The possible biological mechanisms of air pollution on birth weight might vary according the time of pregnancy, such as the implantation of the fetus and the formation of placenta during the first trimester as well as important weight gain during the third trimester. A study found that placental abnormalities due to exposure in the first trimester and complex vascular alterations in the second and third trimester could be the main causes of placental abnormalities and fetal growth retardation(31). PM₁₀ contains many toxic elements including polycyclic aromatic hydrocarbons (PAHs), which could produce DNA adducts and cause DNA damage, resulting in activation of apoptotic pathways(32). Molecular epidemiological studies have been shown that the levels of DNA adducts are positively related to the risk of low birth weight(33-36). Another possible way is that toxic components and/or its metabolites may bind some receptors, which results in disruption of endocrine system and changes the placental function with decreasing exchange of oxygen and nutrients(37). Moreover, these toxic elements might also be able to provoke alveolar inflammation and release the mediators capable of increasing blood coagulability which increase blood viscosity (38). The change of viscosity can affect blood perfusion and have an adverse effect on placental functions, which may result in adverse birth outcomes. All these changes could result into consequent fetal growth retardation. Therefore, our finding of a significant effect during the first trimester is coherent with some potential mechanisms.

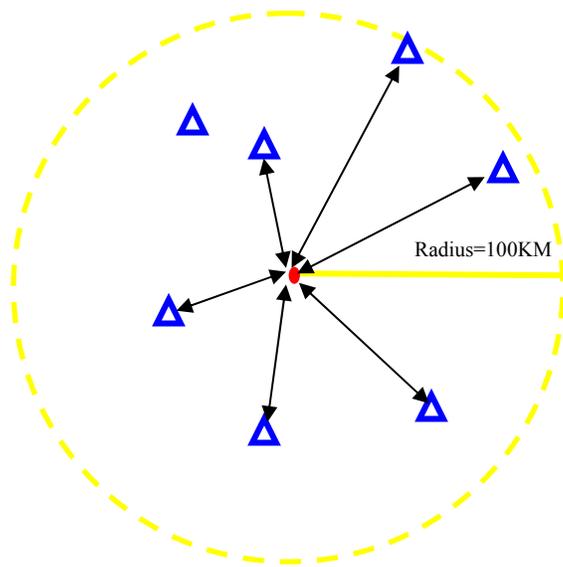
There are number of strengths to our study. First, the present study had a population based birth data with a large sample size to assess the effects of particulate air pollution on term LBW. By focusing on term LBW, the effect of PM₁₀ on fetal growth independent from the effects of prematurity could be examined in this study. Second, the health effects of particulate air

pollutant on birth weight were estimated after controlling for potential confounding factors in this study. Important known risk factors for a low birth weight baby like maternal smoking and maternal weight gain were also considered in this study. Most previous studies of adverse pregnancy outcomes did not control for the information due to unavailability (3-5, 26, 27, 29). Third, our study used the inverse-distance interpolation to predict the levels of particulate air pollutant in smaller geographical resolutions, which provided a more accurate exposure assessment for individual mothers. The inverse-distance interpolation, one of the popular methods to predict spatial distribution of air pollutants, uses real pollution measurements in the neighborhoods and estimates the weight average of neighboring values for un-sampled locations(39).

While the findings of this study are in accordance with other previous studies, some methodological aspects of this study should be elucidated. A certain degree of inaccuracy of maternal exposure might occur due to invalid assumption, difference between monthly period and gestational age, and difference between personal exposure and estimated quarterly measurement. For example, one assumption of this study is that mothers did not move and stayed most of time pregnancy in the place where they resided. However, we do not have information on maternal mobility during pregnancy to assess whether the assumption is true or not. In addition, although our analyses controlled for a number of important potential confounders, information on other factors such as second hand smoking, maternal occupational exposures and nutrition are not available in this study. Therefore, the potential confounding of these factors could not be controlled. However, we adjusted for maternal weight gain in our analyses, which could partially control for the effects of maternal nutrition. Another issue in this study is that we used quarterly levels of PM_{10} to estimate maternal trimester-specific exposure.

Therefore, a certain degree of inaccuracy of maternal exposure might occur. However, such bias or inaccuracy would be non-differential and make the effects toward the null. Moreover, we applied the weight method to estimate the accumulated exposure, which could provide estimates very close to one based on monthly air data.

In conclusion, our findings suggested that exposure to PM10 during the first trimester of pregnancy is associated with an increased risk of term LBW among infants whose mothers resided in Allegheny County, PA, between 1994 and 2000. Further studies are required to elucidate this association and to corroborate the findings of potential public health significance.



▲ Monitor sites
 Points to be

↔ Distance

$$\hat{Z}_{0IDp} = \frac{\sum_{i=1}^N Z(s_i) d_{0,i}^{-p}}{\sum d_{0,i}^{-p}}$$

\hat{Z}_{0IDp} estimated value

$Z(s_i)$ observed value in "i" monitor site

N: number of monitor sites

$d_{0,i}$ distance between estimated point and "i" site

p the power

Figure 3-1 Inverse-distance interpolation

Table 3-1 Weights for estimating the accumulated exposure during pregnancy

Month of conception	Accumulated Exposure first 3 months		Accumulated Exposure first 6 months			Accumulated Exposure first 9 months			
	Q _c	Q _{a1}	Q _c	Q _{a1}	Q _{a2}	Q _c	Q _{a1}	Q _{a2}	Q _{a3}
1	3	-	3	3	-	3	3	3	-
2	2	1	2	3	1	2	3	3	1
3	1	2	1	3	2	1	3	3	2
4	3	-	3	3	-	3	3	3	-
5	2	1	2	3	1	2	3	3	1
6	1	2	1	3	2	1	3	3	2
7	3	-	3	3	-	3	3	3	-
8	2	1	2	3	1	2	3	3	1
9	1	2	1	3	2	1	3	3	2
10	3	-	3	3	-	3	3	3	-
11	2	1	2	3	1	2	3	3	1
12	1	2	1	3	2	1	3	3	2

Q_c : the level of PM10 during the quarter of birth conception

Q_{a1} : the level of PM10 in the 1st quarter after the quarter of birth conception

Q_{a2} : the level of PM10 in the 2nd quarter after the quarter of birth conception

Q_{a3} : the level of PM10 in the 3rd quarter after the quarter of birth conception

Table 3-2 Distribution of trimester-specific PM₁₀, Allegheny County PA, 1994-2000

Trimester (monthly average)	Mean	25%	50%	75%	95%
1 st Trimester	28.7	25.1	28.3	32.2	37.4
2 nd Trimester	28.2	24.9	27.6	31.3	36.9
3 rd Trimester	28.1	24.9	27.3	31.0	36.6
9-month period	28.1	25.8	27.8	30.1	33.5

Table 3-3 Distribution of characteristics of term singleton births by low birth weight status

Characteristics		Term LBW		OR (95%CI)
		CASE (N=2,058)	Control (N=90,389)	
Maternal age				
	<20	307 (14.9)	7,620 (8.4)	1.69 (1.49-1.93)
	20-29*	949 (46.1)	39,907 (44.1)	1.0
	30-34	463 (22.5)	27,628 (30.6)	0.71 (0.63-0.79)
	35-39	282 (13.7)	13,002 (14.4)	0.91 (0.80-1.04)
	≥40	57 (2.8)	2,232 (2.5)	1.07 (0.82-1.41)
Maternal race				
	White*	1,250 (60.9)	72,087 (79.8)	1.0
	Black	731 (35.6)	15,808 (17.5)	2.67 (2.43-2.93)
	Other	73 (3.5)	2,392 (2.7)	1.76 (1.38-2.24)
Maternal education				
	<9	43 (2.1)	641 (0.7)	4.51 (3.27-6.22)
	9-11	313 (15.2)	6,662 (7.4)	3.16 (2.74-3.64)
	12	761 (37.0)	27,512 (30.4)	1.86 (1.66-2.08)
	13-15	425 (20.6)	20,884 (23.1)	1.37 (1.20-1.56)
	≥16 *	516 (25.1)	34,690 (38.4)	1.0
Maternal tobacco use				
	Yes	879 (43.1)	16,103 (17.9)	3.48 (3.18-3.80)
	No	1,160 (56.9)	73892 (82.1)	1.0
Infant sex				
	Male	814 (39.6)	46,407(51.3)	0.62 (0.57-0.68)
	Female *	1,244 (60.45)	43,982 (48.7)	1.0
Parity				
	Second or subsequent birth	1,005 (48.9)	52,856 (58.5)	0.68 (0.62-0.74)
	First birth *	1,049 (51.1)	37,465 (41.5)	1.0
Birth season				
	Summer (Jun-Aug)*	534 (25.9)	23,747 (26.3)	1.0
	Autumn (Sep-Nov)	541 (26.3)	22,155 (24.5)	1.09 (0.96-1.23)
	Winter (Dec-Feb)	503 (24.4)	21,128 (23.4)	1.06 (0.94-1.20)
	Spring (Mar-May)	480 (23.3)	23,359 (25.8)	0.91 (0.81-1.03)
Previous LBW or PTD				
	Yes	53 (2.6)	364 (0.4)	6.56 (4.90-8.77)
	No	1,998 (97.4)	89,931 (99.6)	1.0
Prenatal care				
	None	45 (2.2)	517 (0.6)	3.95 (2.91-5.38)
	1 st trimester *	1,945 (97.3)	88,347 (99.1)	1.0
	After 1 st trimester	10 (0.5)	235 (0.3)	1.93 (1.02-3.64)
Gestational age	Mean±SD	38.2±1.1	39.4±1.2	0.40 (0.39-0.42)
Maternal weight gain	Mean±SD	26.4±11.9	31.0±12.2	0.968 (0.96-0.97)

* Reference group

** Number of missing for each variable: maternal race (106); maternal tobacco use (413); Parity (72); Previous LBW or PTD (101); Prenatal care (1348)

Table 3-4 Risks for term low birth weight ^a according to trimester-specific exposure to PM₁₀

Exposure period	OR (95%CI) *	
	OR	95% CI
1 st Trimester	1.13 ^b	1.02-1.25
2 nd Trimester	1.10	1.00-1.22
3 rd Trimester	1.05	0.96-1.16
9-month period	1.07	0.99-1.14

* ORs were estimated by per inter-quartile range increase (per 7 µg/m³ for trimester-specific exposure and 4.3 µg/m³ for 9-month period) after adjustment for maternal age, maternal race, maternal education, smoking, weight gain, gender of infant, gestation age, parity, previous LBW or preterm birth, level of prenatal care and birth season

^a All singleton births with gestational age ≥37 weeks

^b P<0.05

3.6 REFERENCES

1. Xu X, Ding H, Wang X. Acute effects of total suspended particles and sulfur dioxides on preterm delivery: a community-based cohort study. *Arch Environ Health* 1995;50:407-15.
2. Tsai SS, Yu HS, Liu CC, Yang CY. Increased incidence of preterm delivery in mothers residing in an industrialized area in Taiwan. *J Toxicol Environ Health A* 2003;66:987-94.
3. Bobak M. Outdoor air pollution, low birth weight, and prematurity. *Environ Health Perspect* 2000;108:173-6.
4. Dugandzic R, Dodds L, Stieb D, Smith-Doiron M. The association between low level exposures to ambient air pollution and term low birth weight: a retrospective cohort study. *Environ Health* 2006;5:3.
5. Maisonet M, Bush TJ, Correa A, Jaakkola JJ. Relation between ambient air pollution and low birth weight in the Northeastern United States. *Environ Health Perspect* 2001;109 Suppl 3:351-6.
6. Dejmek J, Selevan SG, Benes I, Solansky I, Sram RJ. Fetal growth and maternal exposure to particulate matter during pregnancy. *Environ Health Perspect* 1999;107:475-80.
7. Liu S, Krewski D, Shi Y, Chen Y, Burnett RT. Association between gaseous ambient air pollutants and adverse pregnancy outcomes in Vancouver, Canada. *Environ Health Perspect* 2003;111:1773-8.
8. Gilboa SM, Mendola P, Olshan AF, et al. Relation between ambient air quality and selected birth defects, seven county study, Texas, 1997-2000. *Am J Epidemiol* 2005;162:238-52.
9. Ritz B, Yu F, Fruin S, Chapa G, Shaw GM, Harris JA. Ambient air pollution and risk of birth defects in Southern California. *Am J Epidemiol* 2002;155:17-25.
10. Moore A. Changing patterns of childhood mortality in Wolverhampton. *Arch Dis Child* 2005;90:687-91.
11. Reynolds P, Urayama KY, Von Behren J, Feusner J. Birth characteristics and hepatoblastoma risk in young children. *Cancer* 2004;100:1070-6.
12. Hales CN. Fetal and infant origins of adult disease. *J Clin Pathol* 1997;50:359.

13. Osmond C, Barker DJ. Fetal, infant, and childhood growth are predictors of coronary heart disease, diabetes, and hypertension in adult men and women. *Environ Health Perspect* 2000;108 Suppl 3:545-53.
14. Morley R. Fetal origins of adult disease. *Semin Fetal Neonatal Med* 2006;11:73-8.
15. Barker DJ. The developmental origins of adult disease. *J Am Coll Nutr* 2004;23:588S-595S.
16. Chen TM, Gokhale J, Shofer S, Kuschner WG. Outdoor Air Pollution: Particulate Matter Health Effects. *Am J Med Sci* 2007;333:235-243.
17. Glorennec P, Monroux F. Health impact assessment of PM10 exposure in the city of Caen, France. *J Toxicol Environ Health A* 2007;70:359-64.
18. Zhang M, Song Y, Cai X. A health-based assessment of particulate air pollution in urban areas of Beijing in 2000-2004. *Sci Total Environ* 2007;376:100-8.
19. Yorifuji T, Yamamoto E, Tsuda T, Kawakami N. Health impact assessment of particulate matter in Tokyo, Japan. *Arch Environ Occup Health* 2005;60:179-85.
20. Chen L, Mengersen K, Tong S. Spatiotemporal relationship between particle air pollution and respiratory emergency hospital admissions in Brisbane, Australia. *Sci Total Environ* 2007;373:57-67.
21. Schwartz J, Slater D, Larson TV, Pierson WE, Koenig JQ. Particulate air pollution and hospital emergency room visits for asthma in Seattle. *Am Rev Respir Dis* 1993;147:826-31.
22. Dockery DW, Pope CA, 3rd, Xu X, et al. An association between air pollution and mortality in six U.S. cities. *N Engl J Med* 1993;329:1753-9.
23. Anderson HR, Spix C, Medina S, et al. Air pollution and daily admissions for chronic obstructive pulmonary disease in 6 European cities: results from the APHEA project. *Eur Respir J* 1997;10:1064-71.
24. Mathews TJ, Hamilton BE. Trend analysis of the sex ratio at birth in the United States. *Natl Vital Stat Rep* 2005;53:1-17.
25. Sram RJ, Binkova B, Dejmek J, Bobak M. Ambient air pollution and pregnancy outcomes: a review of the literature. *Environ Health Perspect* 2005;113:375-82.
26. Ha EH, Hong YC, Lee BE, Woo BH, Schwartz J, Christiani DC. Is air pollution a risk factor for low birth weight in Seoul? *Epidemiology* 2001;12:643-8.
27. Gouveia N, Bremner SA, Novaes HM. Association between ambient air pollution and birth weight in Sao Paulo, Brazil. *J Epidemiol Community Health* 2004;58:11-7.

28. Chen L, Yang W, Jennison BL, Goodrich A, Omaye ST. Air pollution and birth weight in northern Nevada, 1991-1999. *Inhal Toxicol* 2002;14:141-57.
29. Wang X, Ding H, Ryan L, Xu X. Association between air pollution and low birth weight: a community-based study. *Environ Health Perspect* 1997;105:514-20.
30. Jedrychowski W, Bendkowska I, Flak E, et al. Estimated risk for altered fetal growth resulting from exposure to fine particles during pregnancy: an epidemiologic prospective cohort study in Poland. *Environ Health Perspect* 2004;112:1398-402.
31. Lin CC, Santolaya-Forgas J. Current concepts of fetal growth restriction: part I. Causes, classification, and pathophysiology. *Obstet Gynecol* 1998;92:1044-55.
32. Nicol CJ, Harrison ML, Laposa RR, Gimelshtein IL, Wells PG. A teratologic suppressor role for p53 in benzo[a]pyrene-treated transgenic p53-deficient mice. *Nat Genet* 1995;10:181-7.
33. Whyatt RM, Bell DA, Jedrychowski W, et al. Polycyclic aromatic hydrocarbon-DNA adducts in human placenta and modulation by CYP1A1 induction and genotype. *Carcinogenesis* 1998;19:1389-92.
34. Whyatt RM, Santella RM, Jedrychowski W, et al. Relationship between ambient air pollution and DNA damage in Polish mothers and newborns. *Environ Health Perspect* 1998;106 Suppl 3:821-6.
35. Perera FP, Jedrychowski W, Rauh V, Whyatt RM. Molecular epidemiologic research on the effects of environmental pollutants on the fetus. *Environ Health Perspect* 1999;107 Suppl 3:451-60.
36. Perera F, Hemminki K, Jedrychowski W, et al. In utero DNA damage from environmental pollution is associated with somatic gene mutation in newborns. *Cancer Epidemiol Biomarkers Prev* 2002;11:1134-7.
37. Dejmek J, Solansky I, Benes I, Lenicek J, Sram RJ. The impact of polycyclic aromatic hydrocarbons and fine particles on pregnancy outcome. *Environ Health Perspect* 2000;108:1159-64.
38. Maisonet M, Correa A, Misra D, Jaakkola JJ. A review of the literature on the effects of ambient air pollution on fetal growth. *Environ Res* 2004;95:106-15.
39. Spokas K, Graff C, Morcet M, Aran C. Implications of the spatial variability of landfill emission rates on geospatial analyses. *Waste Manag* 2003;23:599-607.

**4.0 CIGARETTE SMOKING AND THE RISK OF ADULT LEUKEMIA: RESULTS
FROM THE THREE MILE ISLAND COHORT STUDY**

Xiaohui Xu, M.P.H.

Evelyn O. Talbott, Dr.P.H., M.P.H.

Jeanne V. Zborowski, M.S., Ph.D.

Judy Rager, M.P.H.

Department of Epidemiology
University of Pittsburgh Graduate School of Public Health

Manuscript in preparation

4.1 ABSTRACT

Smoking is an unconfirmed risk factor for the development of leukemia. The potential link was examined using data from the TMI cohort for the period of 1979 to 1995. A total of 24,539 subjects followed up over 16 years from the TMI cohort who were 14 years or older, were eligible for this analysis. All incident leukemia cases were identified through the Pennsylvania Department of Health Cancer Registry. The Cox proportional hazards model was employed to evaluate the relationships. A total of 42 incident leukemia cases, including 15 AML cases, were observed in the cohort. After controlling for other confounding factors, current smoking was associated with an increased risk of adult AML (RR=3.47; 95% CI: 1.002-11.99). A marginally significant linear trend of risk of AML associated with the number of years smoked was also observed ($p=0.06$). The results from this study suggested that cigarette smoking was associated with an increased risk of adult AML. Further investigation is required to confirm these findings.

Key words: Leukemia, AML, Smoking, Cohort Study.

4.2 INTRODUCTION

Benzene, a well-established carcinogen and leukemogen, has typically been associated with acute myeloid leukemia (AML) over the past century^{1,2} and has been shown to be present in cigarette smoke. Wallace observed that approximately 90% of personal exposure to benzene in the United States is due to smoking. Smokers have an average benzene body burden approximately 6-10 times that of nonsmokers³. Other carcinogens including ionizing radiation (polonium and lead), nitrosamines, styrene, naphthalene and urethanes are also found in cigarette smoking^{4,5}. Cigarette smoking has long been suspected as an etiologic factor in leukemogenesis⁶; an association between cigarette smoking and leukemia was reported as early as 1978⁷. Epidemiological studies to date have provided inconsistent results regarding the association between cigarette smoking and leukemia. Several case-control studies demonstrated an increased risk of leukemia among cigarette smokers⁸⁻¹³. However, other studies have observed no significant association^{14,15}.

Only a few cohort studies have investigated the smoking-leukemia link because of the latency period associated with the development of hematological malignancies and the large sample size required. The conclusions from available prospective studies are inconsistent and some large cohort studies were unable to detect any excess risk of leukemia related to cigarette smoking. Adami et al reported that there was no significant association between smoking status, number of cigarettes smoked per day or duration of smoking and the risk of developing leukemias in 334,957 Swedish construction workers after following up 10 years¹⁶. Although the study is large, all participants are male construction workers. The healthy worker effect cannot be negligible in this study. Engeland et al also found no association between smoking and leukemia in 28 year follow-up of 26,000 Norwegian men and women¹⁷. However, this study

failed to elucidate subtype association and the dose-response relationship of cigarette smoking. Moreover, this study was unable to control for some confounding factors such occupational radiation exposures. In summary, few cohort studies have been carried out to assess the association between cigarette smoking and subtypes of leukemia such as AML, which is more related to the exposure of benzene.

While published evidence for leukemia and cigarette smoking is inconclusive, several case-control studies suggest a weak association between AML and tobacco use^{9,18,19}. The subtypes of leukemia vary in histological, molecular and clinical characteristics as well as in prognosis. The emphasis of etiologic studies on histological subtypes must be taken into consideration.

To further explore the role of cigarette smoking in the development of adult leukemia, we analyzed the risk of total leukemia and specifically acute myeloid leukemia (AML) as well as the dose-response relationship between cigarette smoking and adult leukemia in the large Three Mile Island (TMI) (Pennsylvania, U.S.A) cohort, which provided detailed information on tobacco use at baseline and the incidence of cancer with long-term follow-up.

4.3 METHODS

4.3.1 Study Population

The TMI cohort was assembled by the Pennsylvania Department of Health (PADoH) to evaluate the adverse health effects of exposure of low-level radiation emitted from the TMI nuclear power plant accident on 28 March 1979. The cohort consisted of 32,135 individuals in the 1979 TMI census¹⁸ who were followed up from 1979 to 1995. The estimated average likely

and maximum gamma doses from the accident were 0.09 mSv (9 mrem) and 0.25 mSv (25 mrem) respectively. These exposures were therefore considered overall minimal (less than the dose of 2-3 chest X-rays). Previous mortality studies in these subjects also found no significant evidence of the association between the level of accident radiation exposure and leukemia mortality risk^{18,19}. In this study, all analyses presented are based on cohort members aged 14 years or older (n= 24,539). Data collected included individual information at baseline on smoking status, demographic characteristics (age, gender and race etc), residential history, background radiation exposure, previous occupational or treatment radiation exposure, and estimated radiation exposure from the TMI nuclear power plant accident.

4.3.2 Smoking Information

Data on personal smoking history was collected in 1979 at baseline. Cigarette smoking in individuals was assessed as never smoked, former smoker and current smoker. Former smoker was defined as ever smoked at least 100 cigarettes. For those who ever or currently smoked, The number of cigarettes smoked per day as well as total number of years smoked was also obtained. Based on this information, a new variable of “pack-year” was created by multiplying total number of years smoked with number of cigarettes smoked per day then divided by 20.

4.3.3 Total Leukemia and AML Subtype Identification

The TMI cohort was followed up from March 1979 through December 1995 to determine vital status and cancer incidence. All cases of adult-onset leukemia (International Classification of Diseases for Oncology, Second Edition, and Code 980-994) and AML (ICD-O-2, code 9861,

9864, 9866, 9867 and 9891) diagnosed between March 1979 through December 1995 were identified through the Pennsylvania Cancer Registry (PCR), which is complete for those individuals who remain within Pennsylvania. Not Otherwise Specified (NOS) leukemia cases (ICD-O-2: 9800 and 9801) are further clarified as AML when subject died from AML (International Classification of Diseases, version 9: 205.0) in the National Death Index. The PAdoH, along with the U.S. Postal Service, annually obtained the current addresses of persons to maintain follow-up of the cohort. The address confirmation was updated through December 1995. We verified that 92.1% of subjects remained in Pennsylvania during the 1979-1995 follow-up period.

4.3.4 Statistical analysis

Comparison of the distribution of demographic and other characteristics among each smoking status category was made using ANOVA for continuous variables and the Chi square test for categorical variables respectively. Total leukemia and AML were analyzed separately. The Cox proportional hazards model was employed to estimate Relative Risk (RR) with 95% Confidence Intervals (CI) for the association of adult leukemia and its subtype AML with cigarette smoking after controlling for potential confounding factors. Smoking status was assessed as a dummy variable in the model. The number of cigarettes smoked per day, number of years smoked and number of pack-years were also fit as ordinal categorical variables to evaluate the dose-response relationship between cigarette smoking and total leukemia as well as AML. All final models were adjusted for age, gender, race, education, background radiation exposure ($<8 \mu\text{R/hr}$ vs. $\geq 8 \mu\text{R/hr}$), occupational exposure, and estimated maxima gamma radiation and likely gamma radiation.

Subjects with missing values were excluded from relevant analyses. The statistical tests were two-sided with a 0.05 significant level. The SAS 8.02 software package (SAS Institute Inc., 100 SAS Campus Drive Cary, NC 27513) was used to analyze the collected data.

4.4 RESULTS

Within the study population, 33% (n=8,083) of the individuals were current smokers, 18% (n=4,330) were ever smokers and 49% (11,913) were non-smokers at baseline. More than 95 percent of subjects in the cohort are white. Table 4-1 shows the distribution of baseline characteristics among each group by smoking status. The results of the ANOVA and χ^2 tests indicated that age, education level, race, occupational radiation exposure were significantly different among smoking groups. The percentages of female are 60% among never smokers, 36% among former smoker and 46% among current smokers respectively. There are 5.4% participants among never smokers, 7.4% among former smokers and 6.8% among current smokers who had experience of occupational radiation. The estimated Maximum and likely gamma exposure emitted from the TMI nuclear power plant accident was slightly higher among nonsmokers than smokers. However, the distributions of background radiation exposure were similar among each smoking group.

Table 4-2 describes the distributions of lymphatic and hematopoietic neoplasm incident cases occurred during study period. A total of 42 leukemia cases including 15 AML cases were observed within follow-up time period in the cohort.

The distributions of total leukemia among each status of cigarette smoking and risks of total leukemia associated with cigarette smoking are presented in table 4-3. We found higher proportion of leukemia cases among the current smoker group as compared with those who had never smoked (table 4-3). Of these, the cumulative incidences of total leukemia are 0.21% in current smokers, 0.14% in former smokers and 0.16% in non-smokers respectively.

To examine the association between dose of cigarette smoking and total leukemia, the age and sex adjusted model showed that a positive association between current cigarette smoking and total leukemia was suggestive but not statistically significant by using a dummy variable of smoking status (RR=1.49, 95% CI: 0.76-2.96 in current smoker). After controlling for other risk factors in the fully adjusted model, current cigarette smoking is still positively but not statistically significantly associated with total leukemia (RR=1.39, 95%CI: 0.69-2.82). Furthermore, a more powerful analysis using the number of cigarettes smoked per day, the number of years smoked and pack years of smoking was performed to provide evidence of an association between cigarette smoking and total leukemia. The results showed that there is no evidence of significant positive associations between cigarette smoking and total leukemia either in the age-sex adjusted model or in the fully adjusted model.

Table 4-4 reported the results of the distribution of the subtype of leukemia, i.e. AML among each cigarette smoking group and risk of AML associated with cigarette smoking. Only 15 AML cases were observed within follow-up time period in the cohort. The age-sex adjusted model indicated a significantly positive association between current cigarette smoking and AML using a dummy variable of smoke status (AML: RR=3.26, 95% CI: 1.05-10.17 in current smoker). After controlling for other risk factors in the fully adjusted model, the association between current cigarette smoking and AML is still significant (RR=3.47, 95% CI: 1.002-11.99).

An association between cigarette smoking and AML was further explored with a more powerful analysis using the number of cigarettes smoked per day, the number of years smoked and pack years of smoking. The age-sex adjusted models showed that those who smoked 11-20 pack years have a higher risk of AML as comparing with non-smokers (RR=4.25, 95%CI: 1.11-16.35) after controlling for other risk factors. In addition, there is a significant dose-response relationship between the number years of cigarette smoking and risk of AML ($p=0.04$) (see Fig 1). However, there is no evidence for a linear trend of risk of AML associated with the number of cigarettes per day as well as the pack years of smoking. We further analyzed the risk of AML associated with cigarette smoking after controlling for some other potential confounders, including age, gender, education level, race, occupational radiation, background radiation, estimated maximum and likely gamma exposures. In these analyses, we found a statistically increased risk of AML among those who smoked 1-10 cigarettes per day or 11-20 pack years (RR=5.02, 95% CI: 1.11-22.64; RR=5.07, 95% CI: 1.22-21.01). More interestingly, a dose-response relationship between the number year of smoking and risk of AML was marginally statistically significant ($p=0.06$). However, we found no evidence of any dose-response relationship between the number cigarettes per day or pack years of smoking and risk of AML.

4.5 DISCUSSION

In this large cohort study, 42 incident leukemia cases occurred during the study period. The Not Otherwise Specified (NOS) leukemia was further determined as its subtypes by scrutinizing PA Cancer Registry for cell type and death certificates for underlying cause of death (UCOD). Five NOS leukemia cases were determined as AML (UCOD: 205.0). A total of 15 AML cases

out of 42 leukemia cases were found in this study. The percentage of AML in leukemia is 35.7%, which is close to general population. The analyses of this study emphasized on AML because it had the largest number of cases in leukemia subtypes and epidemiological evidence suggested that AML is more likely linked to cigarette smoking.

In this study, the age-sex adjusted model and the fully adjusted model were applied to assess the associations between leukemia as well as AML and cigarette smoking. The estimations from these two models are very similar. It suggested that age and sex are the factors which confound the associations between smoking and leukemia as well as AML. In this study, neither the age-sex adjusted models nor the fully adjusted models could provide any evidence of a significantly positive association between cigarette smoking and total leukemia after adjusting for other confounders. Indeed, significant association between current cigarette smoking and AML is demonstrated after controlling for other risk factors. It also suggested that there is a significant dose-response relationship between the number year of cigarette smoking and risk of AML (trend $p=0.04$) after adjusted for age and sex. This linear relationship between the number years of cigarette smoking and risk of AML are still marginally significant ($p=0.06$) after controlling for other risk factors. However, there are no statistically significant linear relationships between AML and cigarettes smoked per day as well as pack years. Cigarettes smoked per day, number years of smoked, and pack years have their own limitations and strengths when they are used to estimate the dose of cigarette smoking. Cigarettes smoked per day have information on daily consumption but does not include information on duration of smoking. Number years of smoked emphasizes information on duration of smoking but neglects daily consumption. Pack years is a comprehensive indicator that combines information on daily consumption with duration of smoking. But inaccurate recalls in either cigarettes smoked per day or number years of smoked

could result into a misclassification of pack years. In this study, it could be possible that the recall of years smoked was more accurate than number of cigarettes smoked per day.

The findings of the present study are in agreement with the results from some other studies. The cohort study of 34,000 Seventh-Day Adventists demonstrated that a higher risk of myeloid leukemia was associated with cigarette smoking (RR=2.24, 95% CI 0.91-5.53)²². Several case-control studies reported an increased risk of AML among cigarette smokers^{6,8,9,13,15,23}. Moreover, the childhood leukemia studies also provided significant evidence of an association between parental smoking and childhood leukemia. Chang et al found that parental preconception smoking was significantly associated with an increased risk of AML (OR=3.84, 95%:1.04-14.17)¹⁰. Ji et al reported that paternal preconception smoking was related to a significantly elevated risk of childhood cancers, particularly acute leukemia and lymphoma²⁴. These data are consistent with a possible mechanism linking cigarette smoking to an increased risk of leukemia in human subject.

On the other hand, a few studies reported that the risk of leukemia induced by cigarette smoking decreases or ceases after smoking cessation. Kane et al showed that the odds ratio was decreased as the number of years 'stopped smoking' increased, falling to 1.0 amongst those who had given up smoking for more than 10 years²⁵. Bjork et al found that the risk of AML declined in 5-10 years after smoking cessation among subjects with smoking history long enough to give a substantially elevated AML risk¹. These studies provided supporting evidence of an association between AML and cigarette smoking.

The number of cigarettes per day, the number of years smoked and pack years of cigarette smoking have been previously used to evaluate the dose-response relationship for leukemia and AML. Our study found no evidence of a trend in risk with number of cigarettes smoked per day

and pack years, which was similar to the majority of studies^{9,14,26}. A number of studies reported that there were significant increasing trends in risk with increasing number of years smoked^{13,22}. These findings were consistent with the results from our study, which provided additional support for an association between cigarette smoking and leukemia.

The mechanism by which cigarette smoking increases risk of leukemia or AML are unknown. Cigarette smoke contained some known or suspected leukemogens, including benzene³, ionizing radiation⁵ and other carcinogens (nitrosamines, styrene, naphthalene and urethanes)^{4,27}. It was reported that the level of trans, trans-muconic acid (t,t-MA), a urinary benzene metabolite, were about 3 times higher than those in nonsmokers²⁸. Other studies found higher level of polonium and lead in tissues from smokers²⁹. In addition, chromosomal defects in the peripheral blood were also observed to be increased in smokers^{30,31}. All these findings suggested a possible leukemogenic effect of cigarette smoking.

There are several strengths to the study. First, the study was considered noteworthy, as it included a large sample size of cohort and a relatively long period of follow-up, i.e. over 16 years complete follow-up. Second, the rate of loss of follow-up is low in this study. Ninety three percent of participants in the cohort were verified for leukemia outcome using the population-based Pennsylvania Cancer Registry system. Third, we assessed the risk of AML among cigarette smokers and further evaluated the dose-response relationship between cigarette smoking and outcomes, as this form of leukemia has more specifically been linked to benzene exposure in occupational and community setting. In addition, personal information about natural background radiation, TMI accidental exposure and previous occupational radiation was collected in the study and was able to be included in the final statistical models.

One of the limitations of our study lies in the fact that the baseline cigarette smoking information was only able to obtain in 1979. Any changes in smoking status during the follow-up (1979-1995) were not available. This could result into some misclassification bias in the analyses based on duration of smoking and pack-years of smoking. Second, an additional 12 years of follow up would be very beneficial and improve the power of the study if the cohort could be followed up until now. However, formal follow up of the cohort ceased in 1995 when the PA Department of Health ceased funding formal follow up through post offices of the cohort. Thus, there would no longer be complete ascertainment of the denominator of the population. Third, the analysis was based on leukemia cases identified through the Pennsylvania Cancer Registry (PCR), 1979-1995. Only individuals who remained in Pennsylvania were included in the case ascertainment. In this study, 1,960 cases (7.9%) left the state and lost to follow-up during the study period. The cigarette smoking status and mortality experience were assessed among these individuals of loss follow-up. The distributions of smoking status among these cases are 34% current smokers, 17% former smokers and 49% non-smokers, which is similar to the distributions of the cohort. Moreover, 179 deaths among 7.9% loss follow-up were identified and only one case that died from myeloid leukemia in current smoking group by scrutinizing death certificate for diagnostic code. This result suggested that the study could potentially underestimate the association between cigarette smoking and leukemia. Fourth, although there was a large cohort in this study, only a few new AML cases (n=15) were identified during the follow-up time period because of the low incidence of AML, i.e. 1 to 15 cases per 100,000 persons in the general population³². It resulted into a wide range of 95% Confidence Interval of hazard ratios. However, we had statistical power to detect the risk difference among cigarette smoking groups. Finally, information regarding other potential risk factors of leukemia, such as

diet and medicine, was not obtained in the study. Thus, their potential confounding effects could not be ruled out.

In conclusion, the present study suggests that cigarette smoking is associated with AML. Moreover, the risk of AML was observed to be increased with the increasing number of years smoked. Further well-designed studies are needed to support the establishment of causal association between cigarette smoking and AML.

Table 4-1 Distribution of baseline characteristics among smoking groups

	Never smoked (n=11,913)	Former smoker (n=4,330)	Current smokier (n=8,083)	P-value
Age ($\bar{X} \pm SD$)	40.60 \pm 20.73	46.55 \pm 17.33	38.43 \pm 15.52	<0.001
Gender				
Male (%)	4,746 (39.84)	2,770 (63.97)	4,382 (54.21)	<0.001
Female (%)	7,167 (60.16)	1,560 (36.03)	3,701 (45.79)	
Education level				
≤ 9 yr (%)	2,019 (17.39)	795 (18.51)	1,233 (15.42)	<0.001
10-12 yr (%)	7,037 (60.62)	2,587 (60.25)	5,615 (70.24)	
> 12 yr (%)	2,553 (21.99)	912 (21.24)	1,146 (14.34)	
Race				
White	11,621 (97.55)	4,227 (97.62)	7,836 (96.94)	0.016
Non-white	292 (2.45)	103 (2.38)	247 (3.06)	
Occupational radiation				
Yes	642 (5.39)	321 (7.41)	555 (6.87)	<0.001
No	11,271 (94.61)	4,009(92.59)	7,528 (93.13)	
Background radiation				
< 8 mrem	9,039 (76.12)	3,331 (77.11)	6,191 (76.85)	0.30
≥ 8 mrem	2,836 (23.88)	989 (22.89)	1,865 (23.15)	
Maximum gamma exposure ($\bar{X} \pm SD$)	25.59 \pm 22.08	24.70 \pm 21.85	24.67 \pm 21.75	0.005
Likely gamma exposure ($\bar{X} \pm SD$)	10.93 \pm 10.40	10.61 \pm 10.14	10.58 \pm 10.25	0.04

Note: χ^2 test for the categorical variables and ANOVA test for the continuous variables

Table 4-2 Distribution of lymphatic and hematopoietic neoplasm incident cases occur during study period (1979-1995)

Types of Cancer	Number	ICD-O-2
Leukemia	42	980-994
Acute myeloid leukemia(AML)*	15	9861, 9864, 9866, 9867 and 9891
Chronic lymphocytic leukemia(CLL)	11	9823
Chronic myeloid leukemia (CML)	7	9863
Other leukemia	9	-
Mutiple Myeloma	12	9732
Other Lymphatic Neoplasm	3	-
Total	57	C42.1

* Leukemia NOS (ICD-O-2: 9800, 9801) cases are clarified with death index (ICD-9:205.0)

Table 4-3 Distribution and risk of total leukemia among each status of cigarette smoking

	Total No.	Leukemia		Adjusted RR ¹ (95% CI)	Adjusted RR ² (95% CI)
		No.	%		
Smoking status					
Never smoked (%)	11,913	19	0.16	1.0	1.0
Former smoker (%)	4,330	6	0.14	0.60 (0.23-1.54)	0.58 (0.22-1.50)
Current smoker (%)	8,083	17	0.21	1.49 (0.76-2.96)	1.39 (0.69-2.82)
Number cigarettes per day					
0	11,913	19	0.16	1.0	1.0
<10	1,749	4	0.23	1.37 (0.46-4.05)	1.42 (0.48-4.20)
10-19	2,843	7	0.25	1.61 (0.67-3.86)	1.62 (0.67-3.93)
>=20	7,821	12	0.15	0.84 (0.39-1.78)	0.73 (0.34-1.60)
P value for trend test				0.79	0.57
Number of years smoked					
0	11,913	19	0.16	1.0	1.0
1-19 yrs	7,608	5	0.07	0.55 (0.20-1.52)	0.56 (0.20-1.55)
20-29 yrs	2,015	8	0.40	1.76 (0.76-4.10)	1.73 (0.74-4.07)
>=30 yrs	2,790	10	0.36	1.29 (0.58-2.88)	1.11 (0.48-2.57)
P value for trend test				0.33	0.53
Pack-Year *					
0	11,913	19	0.16	1.0	1.0
≤10	5,897	7	0.12	0.998 (0.41-2.43)	1.02 (0.42-2.50)
11-20	2,358	7	0.30	1.74 (0.72-4.22)	1.74 (0.71-4.27)
>20	4,158	9	0.22	0.88(0.39-2.0)	0.73 (0.31-1.74)
P value for trend test				0.98	0.73

Note: * "Pack-year" is 20 cigarettes per day for one year.

1. "Adjusted RRs" were derived from Cox proportional hazards model after controlling for age and gender.

2. "Adjusted RRs" were derived from Cox proportional hazards model, adjusting for age, gender, education level, race, occupational radiation, background radiation, estimated maximum and likely gamma exposures.

Table 4-4 Distribution and risk of AML among each status of cigarette smoking

	Total No.	AML		Adjusted RR ¹ (95% CI)	Adjusted RR ² (95% CI)
		No	%		
Smoking status					
Never smoked (%)	11,913	5	0.04	1.0	1.0
Former smoker (%)	4,330	1	0.02	0.48 (0.05-4.20)	0.54 (0.06-4.98)
Current smoker (%)	8,083	9	0.11	3.26 (1.05-10.17)	3.47 (1.002-11.99)
Number cigarettes per day					
0	11,913	5	0.04	1.0	1.0
<10	1,749	3	0.17	4.13 (0.98-17.42)	5.02 (1.11-22.64)
10-19	2,843	3	0.11	2.77 (0.65-11.75)	3.31 (0.73-14.99)
>=20	7,821	4	0.05	1.26 (0.32-4.90)	1.07 (0.23-5.01)
P value for trend test				0.67	0.82
Number years of smoking					
0	11,913	5	0.04	1.0	1.0
1-19 yrs	7,608	2	0.03	0.79 (0.15-4.27)	0.91 (0.16-5.19)
20-29 yrs	2,015	3	0.15	3.03 (0.71-12.98)	3.79 (0.82-17.61)
>=30 yrs	2,790	5	0.18	3.58 (0.95-13.50)	3.64 (0.80-16.49)
P value for trend test				0.04	0.06
Pack-Year*					
0	11,913	5	0.04	1.0	1.0
≤10	5,897	3	0.05	1.67 (0.38-7.27)	1.95 (0.42-9.03)
11-20	2,358	4	0.17	4.25 (1.11-16.35)	5.07 (1.22-21.01)
>20	4,158	3	0.07	1.44 (0.33-6.30)	1.11 (0.19-6.36)
P value for trend test				0.32	0.45

Note: * "Pack-year" is 20 cigarettes per day for one year.

1. "Adjusted RRs" were derived from Cox proportional hazards model after controlling for age and gender.
2. "Adjusted RRs" were derived from Cox proportional hazards model, adjusting for age, gender, education level, race, occupational radiation, background radiation, estimated maximum and likely gamma exposures.

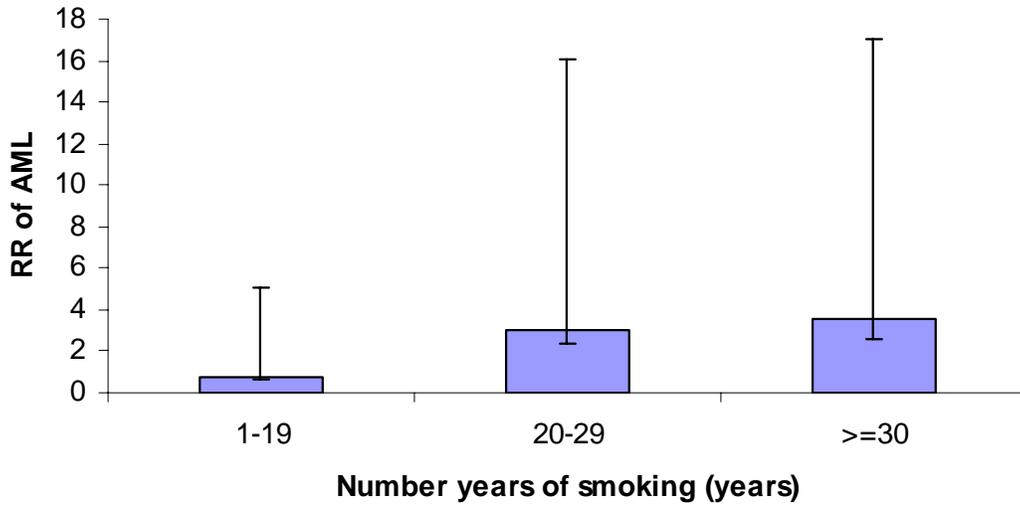


Figure 4-1 the association between risk of AML and the number years of smoking

4.6 REFERENCES

1. Albin M, Bjork J, Welinder H, Tinnerberg H, Mauritzson N, Johansson B, Billstrom R, Stromberg U, Mikoczy Z, Ahlgren T, Nilsson PG, Mitelman F, Hagmar L. Acute myeloid leukemia and clonal chromosome aberrations in relation to past exposure to organic solvents. *Scand J Work Environ Health* 2000;26(6):482-91.
2. Rangan U, Snyder R. Scientific update on benzene. *Ann N Y Acad Sci* 1997;837:105-13.
3. Wallace L. Environmental exposure to benzene: an update. *Environ Health Perspect* 1996;104 Suppl 6:1129-36.
4. Austin H, Cole P. Cigarette smoking and leukemia. *J Chronic Dis* 1986;39(6):417-21.
5. Paffenbarger RS, Jr., Brand RJ, Sholtz RI, Jung DL. Energy expenditure, cigarette smoking, and blood pressure level as related to death from specific diseases. *Am J Epidemiol* 1978;108(1):12-8.
6. Brown LM, Gibson R, Blair A, Burmeister LF, Schuman LM, Cantor KP, Fraumeni JF, Jr. Smoking and risk of leukemia. *Am J Epidemiol* 1992;135(7):763-8.
7. Brownson RC, Chang JC, Davis JR. Cigarette smoking and risk of adult leukemia. *Am J Epidemiol* 1991;134(9):938-41.
8. Chang JS, Selvin S, Metayer C, Crouse V, Golembesky A, Buffler PA. Parental smoking and the risk of childhood leukemia. *Am J Epidemiol* 2006;163(12):1091-100.
9. Pogoda JM, Preston-Martin S, Nichols PW, Ross RK. Smoking and risk of acute myeloid leukemia: results from a Los Angeles County case-control study. *Am J Epidemiol* 2002;155(6):546-53.
10. Severson RK, Buckley JD, Woods WG, Benjamin D, Robison LL. Cigarette smoking and alcohol consumption by parents of children with acute myeloid leukemia: an analysis within morphological subgroups--a report from the Childrens Cancer Group. *Cancer Epidemiol Biomarkers Prev* 1993;2(5):433-9.
11. Severson RK, Davis S, Heuser L, Daling JR, Thomas DB. Cigarette smoking and acute nonlymphocytic leukemia. *Am J Epidemiol* 1990;132(3):418-22.
12. Kabat GC, Augustine A, Hebert JR. Smoking and adult leukemia: a case-control study. *J Clin Epidemiol* 1988;41(9):907-14.

13. Spitz MR, Fueger JJ, Newell GR, Keating MJ. Leukemia and cigarette smoking. *Cancer Causes Control* 1990;1(2):195-6.
14. Adami J, Nyren O, Bergstrom R, Ekblom A, Engholm G, Englund A, Glimelius B. Smoking and the risk of leukemia, lymphoma, and multiple myeloma (Sweden). *Cancer Causes Control* 1998;9(1):49-56.
15. Engeland A, Andersen A, Haldorsen T, Tretli S. Smoking habits and risk of cancers other than lung cancer: 28 years' follow-up of 26,000 Norwegian men and women. *Cancer Causes Control* 1996;7(5):497-506.
16. Pasqualetti P, Festuccia V, Acitelli P, Collacciani A, Giusti A, Casale R. Tobacco smoking and risk of haematological malignancies in adults: a case-control study. *Br J Haematol* 1997;97(3):659-62.
17. Speer SA, Semenza JC, Kurosaki T, Anton-Culver H. Risk factors for acute myeloid leukemia and multiple myeloma: a combination of GIS and case-control studies. *J Environ Health* 2002;64(7):9-16; quiz 35-6.
18. Talbott EO, Youk AO, McHugh-Pemu KP, Zborowski JV. Long-term follow-up of the residents of the Three Mile Island accident area: 1979-1998. *Environ Health Perspect* 2003;111(3):341-8.
19. Talbott EO, Youk AO, McHugh KP, Shire JD, Zhang A, Murphy BP, Engberg RA. Mortality among the residents of the Three Mile Island accident area: 1979-1992. *Environ Health Perspect* 2000;108(6):545-52.
20. Mills PK, Newell GR, Beeson WL, Fraser GE, Phillips RL. History of cigarette smoking and risk of leukemia and myeloma: results from the Adventist health study. *J Natl Cancer Inst* 1990;82(23):1832-6.
21. Kasim K, Levallois P, Abdous B, Auger P, Johnson KC. Lifestyle factors and the risk of adult leukemia in Canada. *Cancer Causes Control* 2005;16(5):489-500.
22. Ji BT, Shu XO, Linet MS, Zheng W, Wacholder S, Gao YT, Ying DM, Jin F. Paternal cigarette smoking and the risk of childhood cancer among offspring of nonsmoking mothers. *J Natl Cancer Inst* 1997;89(3):238-44.
23. Kane EV, Roman E, Cartwright R, Parker J, Morgan G. Tobacco and the risk of acute leukaemia in adults. *Br J Cancer* 1999;81(7):1228-33.
24. Brownson RC. Cigarette smoking and risk of leukemia. *J Clin Epidemiol* 1989;42(10):1025-8.
25. Tso TC, Harley N, Alexander LT. Source of lead-210 and polonium-210 in tobacco. *Science* 1966;153(738):880-2.

26. Hecht SS, Hoffmann D. Tobacco-specific nitrosamines, an important group of carcinogens in tobacco and tobacco smoke. *Carcinogenesis* 1988;9(6):875-84.
27. Obe G, Vogt HJ, Madle S, Fahning A, Heller WD. Double-blind study on the effect of cigarette smoking on the chromosomes of human peripheral blood lymphocytes in vivo. *Mutat Res* 1982;92(1-2):309-19.
28. Melikian AA, Prahalad AK, Hoffmann D. Urinary trans,trans-muconic acid as an indicator of exposure to benzene in cigarette smokers. *Cancer Epidemiol Biomarkers Prev* 1993;2(1):47-51.
29. Hill CR. Polonium-210 in man. *Nature* 1965;208(5009):423-8.
30. DeMarini DM. Genotoxicity of tobacco smoke and tobacco smoke condensate. *Mutat Res* 1983;114(1):59-89.
31. DeMarini DM. Genotoxicity of tobacco smoke and tobacco smoke condensate: a review. *Mutat Res* 2004;567(2-3):447-74.
32. Xie Y, Davies SM, Xiang Y, Robison LL, Ross JA. Trends in leukemia incidence and survival in the United States (1973-1998). *Cancer* 2003;97(9):2229-35.

5.0 SUMMARY AND CONCLUSIONS

The project has been developed to explore the associations between environmental exposures and chronic diseases based on existing data. It has been segmented into three complementary topics as following:

1. Risk of cardiopulmonary diseases with exposure to air pollution
2. Risks of term LBW with exposure to particulate air pollution
3. Risk of leukemia and cigarette smoking

5.1 ASSOCIATION OF RISK OF CARDIOPULMONARY DISEASE AND AIR POLLUTION

A link between respiratory disease and air pollution has been well established in earlier air pollution studies (Holland et al. 1979; Logan 1953; Logan 1956). Recent short-term and long-term epidemiological studies across the world further showed that air pollution exposure is associated with increase risks of cardiopulmonary mortality and morbidity (Daggett et al. 2000; Dockery et al. 1993; Gehring et al. 2006; Katsouyanni et al. 1997; Krewski et al. 2005; Kunzli and Tager 2000; Nyberg and Pershagen 2000; Pope et al. 1995; Pope et al. 2002; Rosenlund et al. 2006; Wietlisbach et al. 1996). Pope et al found that each 10 mg/m³ elevation in fine particulate air pollution was associated with approximately a 6% increased risk of

cardiopulmonary mortality (Pope et al. 2002). Cohen and his colleagues conducted a risk assessment of the global burden of disease attributable to urban ambient air pollution and showed that ambient air pollution, in terms of fine particulate air pollution, causes about 3% mortality from cardiopulmonary disease and about 1% mortality from acute respiratory infections in children under 5 years worldwide (Cohen et al. 2005).

In this study, we also found that exposure to particulate matter is associated with an increased risk of cardiopulmonary diseases before the LTV coke plant closure, which is consistent to the findings of previous studies.

Studies of cigarette smoking found that the cessation of smoking could result into reduction of health risks (Andrews and Tingen 2006; Bosetti et al. 2006; Krall et al. 2006). However, few air pollution studies have been conducted to assess how elimination or addition of a major point source of air pollution in an area affects the health among people who live there. In this study, we found that the elimination of a major point source of air pollution could reduce health risks in the area. This reduction of risk could possibly be explained by the declination of air pollution concentration and the change of components of particulate matter after the removal of the major point source of air pollution.

The studies across the world provided consistent and coherent toxicological and epidemiological evidence indicating that air pollution exposure is associated with cardiopulmonary disease. Further study might be focused on elucidating the biological mechanism, which is not yet clear.

5.2 ASSOCIATION OF RISK OF TERM LBW AND PARTICULATE AIR POLLUTION

Study of particulate matter on fetal health is an emerging area. Fetuses, a susceptible subgroup of population, are thought to be vulnerable to the effects of air pollution. Birth weight is one of the widely-used adverse pregnancy outcomes in fetal health air pollution studies. The findings of previous studies are not consistent and provide a mixed picture about the association between particulate matter and birth weight. Some studies reported a negative association over some exposure period (Ha et al. 2001; Lee et al. 2003; Mannes et al. 2005; Parker et al. 2005; Rogers and Dunlop 2006; Wang et al. 1997; Wilhelm and Ritz 2005; Yang et al. 2003) and others reported no evidence of an association between particulate matter and low birth weight (Chen et al. 2002; Dugandzic et al. 2006; Maisonet et al. 2001).

Birth weight is sensitive to many factors such as gestational age, maternal age, race, education, maternal weight gain during pregnancy, prenatal care level, smoking, infant sex, parity and others. Without adequate control for confounders, the association between low-level particulate matter and low birth weight might be difficult to quantify. In this study, we found that the odds ratios of term LBW per inter-quartile range increase in PM_{10} were 1.13 (95%CI: 1.02-1.25) during the first trimester and 1.10 (95%CI: 1.00-1.22) during the second trimester after adjusting for other important covariates, respectively.

Further studies with more refined methodological designs such as high-quality exposure data, advanced methods of exposure assessment, and high-quality covariate information, are needed to clarify the adverse effect of particulate air pollution on fetal health.

5.3 ASSOCIATION OF RISK OF LEUKEMIA AND CIGARETTE SMOKING

Cigarette smoking is associated with many health outcomes such as chronic obstruction pulmonary disease (Davis and Novotny 1989; Kalucka 2006; Lee et al. 1990), cardiovascular disease (Cook et al. 1986; Lapidus et al. 1986; Milei et al. 1986), lung cancer (Benhamou et al. 1985; Kestner 1985; La Vecchia 1985), nasopharyngeal cancer (Mabuchi et al. 1985), bladder cancer (Marrett et al. 1985; Rebelakos et al. 1985), and adverse pregnancy outcomes (Weisberg 1985). Cigarette smoking has not classically been linked with leukemia until a few epidemiological studies reported an increased risk of leukemia among cigarette smokers (Brownson 1989; McLaughlin et al. 1989; Severson et al. 1990; Spitz et al. 1990). Several case-control studies suggested an association that is most pronounced for acute myeloid leukemia (AML) and provided evidence of a dose-response relationship between cigarette smoking and AML (Chelghoum et al. 2002; Kasim et al. 2005; Sandler et al. 1993).

In this study, we used a population based cohort to evaluate the association between leukemia as well as AML and cigarette smoking. The results showed that cigarette smoking is significantly associated with AML after controlling for other important risk factors. In addition, a dose-response relationship between cigarette smoking and AML, which represents additional evidence that cigarette smoking is an etiological factor of AML, was also observed. The finding in this cohort study is consistent with the results of previous case-control studies.

One of the strengths of a cohort study, as compared with a case-control study, is that the true risk of disease can be estimated in the study. A major limitation of cohort study for rare disease like AML is that the study sample size should be very large to detect significant associations for rare diseases like AML. Therefore, a few AML cases observed during the study period could

limit the value of this study. This might explain the wide range of risk estimation although there is not small population with 16 year follow-up in the study.

The epidemiological evidence concerning the possible etiologic role of cigarette smoking in leukemia as well as AML are few and controversial. A well designed cohort study could be more valuable than case-control study in establishing a relationship of causality between cigarette smoking and leukemia. Further studies could be improved by using less biased information of cigarette smoking, high quality covariate data, accurate classification of leukemia subtypes and a larger study sample size.

The findings of studies of cigarette-related leukemia could reinforce tobacco control including preventing initiation, promoting cessation, reducing exposure to cigarette smoking and finally reducing the health burden of cigarette-related diseases in the population.

5.4 STRENGTH AND LIMITATIONS OF SECONDARY DATA

Environmental public health tracking might cover a large population with a long time period in a region. Primary data gathering for environmental health tracking is unrealistic and inefficient. A variety of secondary data including environmental hazard, exposure and health outcomes is readily available in federal, state and local environmental and health agencies. It is a great source for environmental public health tracking. The review of three types of data including environmental hazard, environmental exposure and health outcomes is discussed in the section of Appendix B. Therefore, secondary data could provide an excellent and economical source for environmental public health tracking in comparison with collecting primary data.

Using secondary data might be an effective and quick way to identify health problems related to environmental exposure in a region.

However, there are a number of limitations when using secondary data for environmental public health tracking. The availability of information and data structure in secondary data heavily influence the study design and data analyses. There is often no way to validate the quality of secondary data because of lack of information about problems of sampling, method of data collection, response rates, data coding and others. We may lack information about the strengths and weaknesses of secondary data which are used in environmental public health tracking. The validity and reliability of results based on these data are difficult to assess.

Although secondary data is readily available, the accuracy and completeness may also vary. Therefore, some data issues including completeness of data, accuracy of information collected, registration period, data format, data accessibility and availability and possibilities of linkage with other data sources should also be aware of when secondary data are used in environmental public health tracking (See Appendix B).

Overall, the limitations of secondary data can not be avoided because it is usually collected for purposes other than for environmental public health tracking. However, secondary data is a valuable source and still attractive to environmental public health tracking agencies.

5.5 OVERALL SUMMARY AND CONCLUSION

The overall objectives of this project is to demonstrate how to using secondary data to link environmental exposures to chronic disease, to evaluate the possible associations between environmental exposure and health outcome and to explore the methodology of study design,

data linkage and data analyses. In this project, we succeeded in linking a variety of health outcome data including hospital admission data, birth registry data and research cohort data, with environmental exposures and examining the potential associations between environmental exposures and chronic diseases with appropriate methods of study design and data analyses. The techniques used in this project could be applicable to state or local environmental health tracking agencies for conducting similar studies.

APPENDIX A

RISK OF ADVERSE PREGNANCY OUTCOMES AND AIR POLLUTION

The study of adverse pregnancy outcomes is an emerging field of environmental epidemiology. Studies of air pollution on adverse pregnancy outcomes have been performed since the late 1990s. Low birth weight (birth weight less than 2500 g), as one of impaired birth outcomes, has been reported to influence the health status of individuals, including increased mortality and morbidity in childhood (1, 2) and an elevated risk of hypertension, coronary heart disease, type II diabetes in adulthood, abnormalities of lipid metabolism and blood coagulation (3, 4). Low birth weight is a heterogeneous outcome which could be caused by early delivery and fetal growth retardation and is associated with many factors. The early delivery due to premature rupture of the membranes and placenta abruption may result into a low birth weight infant. Some other maternal prenatal factors such as maternal diseases, maternal weight gain and maternal smoking could also be associated with a low birth weight baby. In addition, malformation may also increase the risk of low birth weight infant (5). There is still a limited literature regarding the effects of air pollution on birth weight. This appendix reviewed previous studies that have linked maternal exposure to ambient air pollution to birth weight.

We searched all publications included in the electronic databases PubMed and Ovid Medline with the combination of key words “air pollution” with any following: “birth weight”, “low birth weight”, “LBW”, “very low birth weight” or “VLBW”. Eighteen studies of assessing the relationship between birth weight and ambient air pollution were identified (6-23). These studies were summarized in Table 2. The air pollutant assessed in the studies included total suspended particles (TSP), PM₁₀, PM_{2.5}, SO₂, O₃, CO, NO₂ and NO_x. The estimations of an association between low birth weight and air pollutants reported in these studies ranged approximately from 1.01 to 1.5 (Table 2). The strengths that these studies shared include trimester-specific exposure estimations which could show the effect periods of air pollution on birth weight during pregnancy, standard definitions of adverse birth outcomes and adjustment for other important risk factors of low birth weight. However, the findings from these studies are inconsistent, especially regarding the effect period and the strength of association of particulate matter. In addition, many of these studies didn't account for important confounders such as maternal smoking, gestational age and weight gain. Previous studies have also had limited spatial information on pollution sources and concentrations because the exposure mostly relied on the measurements at monitoring stations. This study, while not capable of clearing up all previous questions, could provide additional evidence regarding the health effects of particulate matter on LBW.

Table A-1 Results of studies of ambient air pollutants and LBW and VLBW

Reference	Design	Pollutants	Outcomes	Results	Covariates/stratified variables
Wang X et al 1997	Cohort study (Beijing, China)	SO ₂ (per 100µg/m ³)	LBW	1.11 (1.06-1.16)	Gestational age, residence, year of birth, maternal age, and infant gender
		TSP(per 100µg/m ³)		1.10 (1.05-1.14)	
Bobak M and Leon DA 1999	Ecological study (Czech Republic)	TSP(per 50µg/m ³)	LBW	1.04 (0.96 -1.12)	socioeconomic factors (mean income, car ownership, divorce rate, etc)
		SO ₂ (per 50µg/m ³)		1.10(1.02- 1.17)	
		NO _x (per 50µg/m ³)		1.07(0.98- 1.16)	
Bobak M 2000	Case-control study (Czech Republic)	TSP(50µg/m ³)	LBW	LBW: 1.15[1.07- 1.24] 1 ST Trimester	Maternal sex, parity, age, education, marital status and month of birth.
		SO ₂ (50µg/m ³)		LBW: 1.20[1.11- 1.30] 1 ST Trimester	
Rogers JF 2000	Case-control study (Georgia, U.S.A)	Sum of TSP and SO ₂ (µg/m ³)	VLBW (<1,500g)	2.88[1.16-7.13] for >95 th percentile 1.27[0.68-2.37] for >75 th to 95 th percentile	Maternal race, age, education, smoking, medicine, weight gain, alcohol, stress and infant sex
Ritz B and Yu F 1999	Cohort study (South California, U.S.A)	CO (>5.5 ppm 3-month average) during the last trimester	LBW	1.22(1.03-1.44)	Commuting habits, infant sex, prenatal care, maternal age, ethnicity, and education

Table A-1 continues

Reference	Design	Pollutants	Outcomes	Results	Covariates/stratified variables
Bobak M 2001	Cross-sectional study (Britain)	Pollution index	Birth weight		Gender, parental social class, maternal education, region, birth order, quality of house
		1 (low)			
		2		3.9(24.8)	
		3		-55.7(26.7)	
		4(high)		-81.4(29.4)	
Maisonet M 2001	Case-control study (Six cities, U.S.A)	CO (1ppm)	LBW	1.31[1.06-1.62]	Maternal age, education, race, smoking, alcohol, marital status, weight gain, previous termination, infant sex, gestational age, season of birth, firstborn, prenatal care.
		SO2			
		25 th to 50 th		1.18[1.12-1.25]	
		50 th to 75 th		1.12[1.07-1.17]	
		75 th to 95 th		1.13[1.05-1.22]	
Ha EH 2001	Case-control study (Seoul, South Korea)	CO	LBW	1.08[1.04-1.12]	Gestational age, maternal age, education, parity, infant's birth order and sex.
		NO2		1.07[1.03-1.11]	
		SO2		1.06[1.02-1.10]	
		TSP		1.04[1.00-1.08]	
		O3		1.09[1.04-1.14]	
Chen L 2002	Cross-sectional study (Northern Nevada, U.S.A)	PM10(per 10µg/m ³)	Reduction of Birth weight	11 [2.3, 19.8]	Maternal age, race, education, smoking, drug, alcohol, prenatal visits, weight gain, residential city, infant sex, gestational age
Maroziene L 2002	Case-control study (Kaunas, Lithuania)	Formaldehyde	LBW	OR _{med.} 1.86 [1.10-3.16]	Maternal age, parity, marital status, education, season of birth and smoking
				OR _{high} 1.84 [1.12-3.03]	
		NO2		OR _{high} 1.64 [1.04-2.58]	

Table A-1 continues

Reference	Design	Pollutants	Outcomes	Results	Covariates/stratified variables
Lee BE 2003	Case-control study (Seoul, Korea)	PM10	LBW	1.03 [1.00-1.07]	Maternal age, education, infant sex, order, gestational age and date
		CO		1.04 [1.01-1.07]	
		PM10		1.04 [1.00-1.08]	
		CO		1.03 [1.00-1.06]	
		SO2		1.06 [1.02-1.11]	
		NO2		1.03 [1.01-1.06]	
Yang CY 2003	Cross-sectional study(Kaohsiung, China)	PM10	Birth weight	β 0.52 [0.19-0.85]	Maternal age, education, marital status, season of delivery and infant sex
		SO2		β 0.52 [0.09-2.63]	
Liu S 2003	Case-control study (Vancouver, Canada)	SO2 (5 ppb)	LBW	1.11 [1.01-1.22]	Maternal age, parity, infant sex, gestational age and season of birth
Gouveia N 2004	Cross-sectional study (Sao Paulo, Brazil)	PM10(per 10 μ g/m ³)	Change of birth weight	-13.7[-27.0, -0.4]	Maternal age, education, parity, antenatal care, gestational age
		CO (per 1ppm)		-23.1 [-41.3, -4.9]	
Parker JD 2005	Cross-sectional study (California, U.S.A)	PM2.5 (per 10 μ g/m ³)	Birth weight	β -29.3g [-42.2, -16.4g]	Maternal race, age, education, parity, marital status, season of delivery and co-pollutant
		CO		β -38.2g [-54.9, -21.6g]	
		PM2.5(per 10 μ g/m ³)	SGA	AOR 1.20[1.07-1.37]	
		CO (per 0.5ppm)	AOR 0.89[0.77-1.03]		

Table A-1 continues

Mannes T 2005	Cohort study (Sydney, Australia)	PM10	Small for gestational age & birth weight	1.02 [.01-1.03]	Infant sex, gestational age, maternal age, race, SES and season birth
		PM2.5		1.03 [.01-1.05]	
		CO		1.10 [0.96-1.27]	
		NO2		1.14 [1.07-1.22]	
		O3		1.01 [1.00-1.01]	
Wilhelm M 2005	Case control study (Los Angeles, CA)	CO	LBW	LBW: 1.36[1.04-1.76] 3 rd trimester	Infant sex, gestational age, maternal age, race, education, prior stillbirth, PTD, parity and birth season, other pollutants
		PM10		LBW: 1.48[1.00-2.19] 3 rd trimester	
Dugandzic R 2006	Retrospective cohort study (Nova Scotia, Canada)	PM10	LBW	RR 1.33 [1.02-1.74] >75 th percentile, 1 st trimester	Gender of infant, gestation age, maternal age, parity. Smoking, weight change, prior neonatal deaths, stillbirth and LBW, and neighborhood family income
		SO2		RR 1.36 [1.04-1.78] >75 th percentile, 1 st trimester	
		O3		RR 1.16 [0.88-1.53] 25 th to 50 th percentile, 1 st trimester	

A.1 REFERENCE

1. Moore A. Changing patterns of childhood mortality in Wolverhampton. *Arch Dis Child* 2005;90:687-91.
2. Reynolds P, Urayama KY, Von Behren J, Feusner J. Birth characteristics and hepatoblastoma risk in young children. *Cancer* 2004;100:1070-6.
3. Hales CN. Fetal and infant origins of adult disease. *J Clin Pathol* 1997;50:359.
4. Osmond C, Barker DJ. Fetal, infant, and childhood growth are predictors of coronary heart disease, diabetes, and hypertension in adult men and women. *Environ Health Perspect* 2000;108 Suppl 3:545-53.
5. Maisonet M, Correa A, Misra D, Jaakkola JJ. A review of the literature on the effects of ambient air pollution on fetal growth. *Environ Res* 2004;95:106-15.
6. Wang X, Ding H, Ryan L, Xu X. Association between air pollution and low birth weight: a community-based study. *Environ Health Perspect* 1997;105:514-20.
7. Bobak M, Leon DA. Pregnancy outcomes and outdoor air pollution: an ecological study in districts of the Czech Republic 1986-8. *Occup Environ Med* 1999;56:539-43.
8. Ritz B, Yu F. The effect of ambient carbon monoxide on low birth weight among children born in southern California between 1989 and 1993. *Environ Health Perspect* 1999;107:17-25.
9. Bobak M. Outdoor air pollution, low birth weight, and prematurity. *Environ Health Perspect* 2000;108:173-6.
10. Rogers JF, Thompson SJ, Addy CL, McKeown RE, Cowen DJ, Decoufle P. Association of very low birth weight with exposures to environmental sulfur dioxide and total suspended particulates. *Am J Epidemiol* 2000;151:602-13.
11. Bobak M, Richards M, Wadsworth M. Air pollution and birth weight in Britain in 1946. *Epidemiology* 2001;12:358-9.
12. Maisonet M, Bush TJ, Correa A, Jaakkola JJ. Relation between ambient air pollution and low birth weight in the Northeastern United States. *Environ Health Perspect* 2001;109 Suppl 3:351-6.
13. Ha EH, Hong YC, Lee BE, Woo BH, Schwartz J, Christiani DC. Is air pollution a risk factor for low birth weight in Seoul? *Epidemiology* 2001;12:643-8.

14. Chen L, Yang W, Jennison BL, Goodrich A, Omaye ST. Air pollution and birth weight in northern Nevada, 1991-1999. *Inhal Toxicol* 2002;14:141-57.
15. Maroziene L, Grazuleviciene R. Maternal exposure to low-level air pollution and pregnancy outcomes: a population-based study. *Environ Health* 2002;1:6.
16. Lee BE, Ha EH, Park HS, et al. Exposure to air pollution during different gestational phases contributes to risks of low birth weight. *Hum Reprod* 2003;18:638-43.
17. Yang CY, Tseng YT, Chang CC. Effects of air pollution on birth weight among children born between 1995 and 1997 in Kaohsiung, Taiwan. *J Toxicol Environ Health A* 2003;66:807-16.
18. Liu S, Krewski D, Shi Y, Chen Y, Burnett RT. Association between gaseous ambient air pollutants and adverse pregnancy outcomes in Vancouver, Canada. *Environ Health Perspect* 2003;111:1773-8.
19. Gouveia N, Bremner SA, Novaes HM. Association between ambient air pollution and birth weight in Sao Paulo, Brazil. *J Epidemiol Community Health* 2004;58:11-7.
20. Parker JD, Woodruff TJ, Basu R, Schoendorf KC. Air pollution and birth weight among term infants in California. *Pediatrics* 2005;115:121-8.
21. Mannes T, Jalaludin B, Morgan G, Lincoln D, Sheppard V, Corbett S. Impact of ambient air pollution on birth weight in Sydney, Australia. *Occup Environ Med* 2005;62:524-30.
22. Wilhelm M, Ritz B. Local variations in CO and particulate air pollution and adverse birth outcomes in Los Angeles County, California, USA. *Environ Health Perspect* 2005;113:1212-21.
23. Dugandzic R, Dodds L, Stieb D, Smith-Doiron M. The association between low level exposures to ambient air pollution and term low birth weight: a retrospective cohort study. *Environ Health* 2006;5:3.

APPENDIX B

USING SECONDARY DATA FOR ENVIRONMENTAL PUBLIC HEALTH

TRACKING: SOURCES, METHODS, AND ISSUES

The environment plays an important role in human health. In the 20th Century, the growing population and industrialization have adversely impacted the natural environment. Many studies have reported an association between chemical, physical and biological environmental exposures and adverse health effects. Adverse health effects related to chemical agents previously were the focus of studies primarily of the occupational environment. However, an enormous number and variety of chemicals have been introduced into the environment in the last decades and communities are now at the core of investigational efforts.

Chemical agents can enter human body through food, drinking water, air and skin contact. There is a considerable concern about the health effects of these agents on the general population. Certain specific environmental hazards have been linked with chronic diseases in the research literature. For instance, epidemiological studies have reported association between chlorinated water consumption and increased risk of various cancers, including cancers of the bladder, rectum, pancreas, kidney, stomach and lymphatic system (1-7). Air pollutant exposures have been associated with increased morbidity and/or mortality of cardiopulmonary disease (8-

11), asthma exacerbation (12-14) and adverse reproductive health outcomes (15-18). Benzene is well recognized as an etiological risk factor for leukemia(19, 20). Toxic effects related to metal or metal compounds have also been identified. For instance, exposure to arsenic has been linked with cancers of the skin(21-24) , lung(25), urinary bladder(26), kidney(25) and liver(27), as well as non-cancer diseases such as peripheral vascular disease(28), cardiovascular and cerebrovascular disease(29), diabetes(30), and adverse reproductive outcomes(31, 32). Lead exposure has been associated with mental retardation(33-35). Physical environmental factors such as noise, extreme weather conditions, low-level radiation exposure and electromagnetic exposure, have also generated great concern in communities. Extremes of temperature has been reported to be associated with short-term increases in daily mortality(36, 37). Residential exposure to radon has been reported to cause lung cancer in human(38, 39). Several studies has linked exposure to electromagnetic fields with increased risk of cancer(40-42). There is little debate that agents in environment could cause adverse health effects in human. A systematic surveillance (e.g. tracking system) for environmental public health is clearly a necessary, rational and appropriate response to community, academic and governmental concerns about environmental exposures.

In 2000, the Pew Environmental Health Commission at the Johns Hopkins School of Hygiene and Public Health recognized that the environmental public health system was fragmented, neglected and ineffective in the United States (43). Under the recommendations of the Pew Environmental Health Commission, the Centers for Disease Control and Prevention (CDC) established the National Environmental Public Health Tracking (EPHT) Program, which is the ongoing collection, integration, analysis, interpretation, and dissemination of data on environmental hazards; exposure to those hazards; and related health effects (44). Its goal is to

provide information that can be used to plan, apply, and evaluate actions to prevent and control environmentally related diseases.

Thacker (1996) presented a useful framework of a “hazard-exposure-outcome” axis for conducting environmental public health tracking. It outlined the steps in the process of how an environmental agent produces adverse health outcomes (see figure B-1)(45). CDC applied the conceptual model proposed by Thacker to design the national EPHT network, which will meld data related to hazards, exposure and health outcomes into a network of standardized, distributed electronic data systems and will provide valid scientific information on environmental exposures and adverse health effects and the possible spatial and temporal relations between them (46).

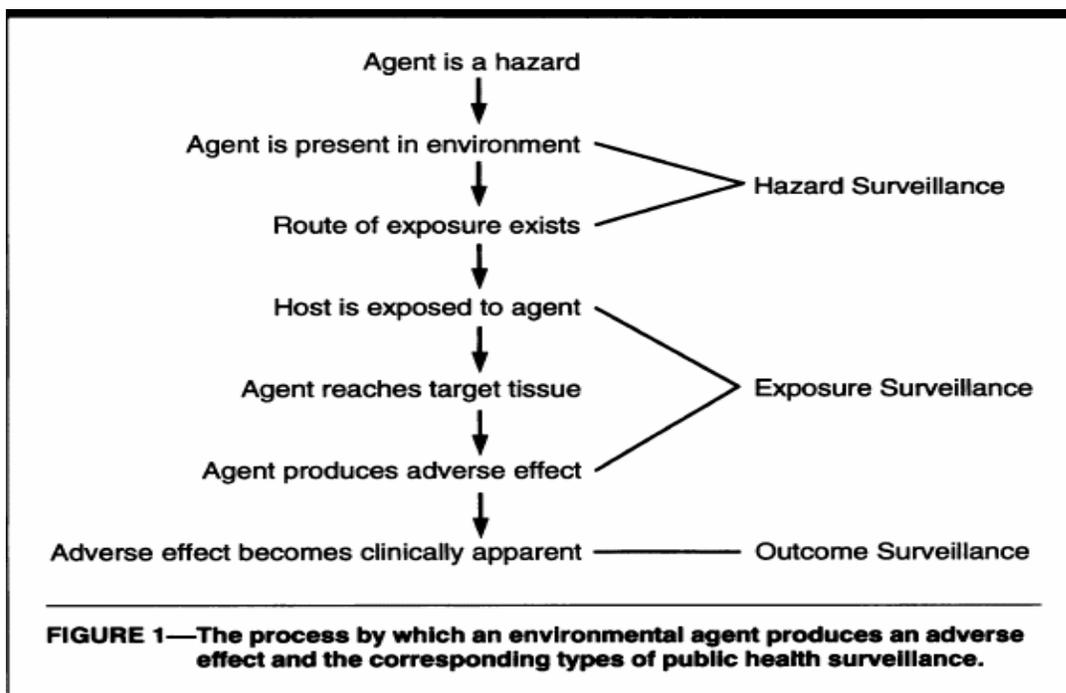


Figure B-1 Environmental agent and adverse health outcomes

(SO: *Am. J. Public Health*, 1996, 86(5): 633-638)

In this review, aspects related to the use of existing data for conducting environmental public health tracking will be discussed, including the availability of existing data, the strategy of data

linkage, statistical framework, issues related to the use of existing data for conducting environmental public health tracking.

B.1 OVERVIEW OF AVAILABLE EXISTING DATA SYSTEMS

The data on hazards, exposures and health outcomes are three types of data in the process by which an agent in the environment can produce adverse health effects in a host. The availability of these data is important for conducting environmental public health surveillance. Data on hazards, exposures, outcomes and spatial data have been collected at the national, regional or local levels as well as in scientific research studies. Much of the data has not been well utilized or never analyzed or interpreted in a way that could address other environmental issues or ever released to the public. A brief description of existing data on environmental hazards, exposures, health outcomes and geospatial field is presented in the following section.

B.1.1 HAZARD DATA

A hazard can be any chemical, physical or biological environmental agent which has a potential for adverse health effects. Information on the presence and quantity of contaminants in environmental media is systematically collected by various organizations. Most hazard data is collected by federal and state agencies and is mandated by legislation, such as Clean Air Act or Clean Water Act (1970). The U.S. Environmental Protection Agency (EPA) was established to consolidate in one agency a variety of federal research, monitoring, standard-setting and

enforcement activities to ensure environmental protection on December 2, 1970. Many hazard databases on air and water have been routinely established and maintained by EPA.

Air

The EPA Air Quality System (AQS) is a database containing hourly and daily measurements of criteria air pollutants and hazardous air pollutants from monitors across the United States. Detailed data from 1994 to the present can be downloaded from the EPA website (<http://www.epa.gov/ttn/airs/airsaqs/detaildata/downloadaqsddata.htm>). The EPA National Emissions Inventory (NEI) database includes estimates of annual pollutant emissions from point, non-point and mobile sources for 50 states, Washington DC, Puerto Rico and the Virgin Islands. Data drawn from the NEI and AQS databases consist of the EPA AirData/AIRS database. These data are available from 1971 to the present.

Toxic Release Inventory (TRI) is an EPA database containing estimated information about the amount of chemicals annually released to air, water, and land by the manufacturing industry. Any facility meeting established criteria (ref for criteria: http://epa.gov/tri/guide_docs/2001/brochure2000.pdf) is required to report annual releases of nearly 600 chemical compounds. The data are available from 1987 to the current time.

Water

EPA maintains two data management systems containing water quality information for the nation's waters: the Legacy Data Center (LDC), and STORET. The LDC is a static, archived database and STORET is an operational system actively being populated with water quality data. The LDC contains historical water quality data dating back to the early part of the 20th century and collected through the end of 1998. STORET contains data collected beginning in 1999. Both databases contain raw biological, chemical, and physical data on surface and ground water

collected by federal, state and local agencies, Indian Tribes, volunteer groups, academics, and others.

EPA also maintains several databases on drinking water. For example, the Safe Drinking Water Information System - Federal version (SDWIS/FED) database includes information about the nation's 160,000 public water systems and violations of drinking water regulations. The National Contaminant Occurrence Database (NCOD) contains occurrence data from both Public Water Systems (PWSs) and other sources (like the U.S. Geological Survey National Water Information System) on physical, chemical, microbial and radiological contaminants for both detections and non-detects. The Unregulated Contaminant Monitoring Rule (UCMR) database is collected to evaluate and prioritize contaminants on the Drinking Water Contaminant Candidate List, which could be possible new drinking water standards. The Federal Reporting Database System (FRDS) is a centralized database for information on public drinking water supplies. It is maintained by the EPA Office of Drinking Water. FRDS contains approximately 12 million records.

In addition, other organizations and institutions maintain nation wide hazard databases. For instance, the U.S. Geological Survey (USGS) created a National Water Information System (NWIS) database which provides access to water-resources data collected at approximately 1.5 million sites in all 50 states, plus border and territorial sites. There are nearly 70 million water-quality results from about 4 million water samples collected at hundreds of thousands of sites. The data are from as early as 1899 to the present.

The USGS implemented the National Water-Quality Assessment (NAWQA) Program in 1991 to develop long-term, consistent and comparable information on streams, rivers, ground

water, and aquatic systems. It includes 162 sample sites in 51 of the nation's river basins and aquifers.

Climate

The National Climate Data Center (NCDC) maintains the world's largest active archive of weather data. The NCDC's database provides a national resource for climate information from 1985 to present. The data include hourly temperature, wind, rain, pressure, clouds and snow in national or global scale.

Agriculture

The National Agricultural Statistics Service (NASS) database is maintained by the United States Department of Agriculture (USDA). The NASS conducts hundreds of surveys every year. It provides information about U.S. agricultural chemical use, agricultural-production and supplies of food and fiber, prices paid and received by farmers, farm labor and wages, farm finances, and changes in the demographics of producers. The available data are from 1867 to the present.

Moreover, some environmental hazard data were created in large national-wide survey studies. For example, the U.S EPA conducted a five-year study entitled National Pesticide Survey (NPS) in 1988. It provided information on specific contaminants, pesticides and nitrates in drinking water from groundwater supplies such as community water system wells and rural domestic wells.

The National Atrazine Occurrence Monitoring Program was conducted under the joint sponsorship of the Association Water Works Association's (AWWA) Water Industry Technical Action Fund (WITAF) and the American Water Works Association Research Foundation

(AwwaRF). It is a study to characterize atrazine occurrence patterns across the United States. Forty-seven drinking water plants were included in this study.

B.1.2 EXPOSURE DATA

Exposure data include bio-monitoring for the presence and the quantity of a compound or its metabolites in human tissues or biologic samples such as fat, blood, hair and urine. Tracking actual exposure to hazards is frequently the missing link in evaluating the risk of environmental hazard. Exposure data are usually monitored at the individual level. However, little individual level exposure data currently exist in accessible datasets.

The National Human Adipose Tissue Survey database is a nation wide exposure data set. It provides information on 54 chemicals in human adipose tissue in three different age groups and in nine different geographic regions in the United States. The database contains results of analyses of human tissue samples collected between 1967 and 1990.

Biomonitoring measurements were also made in samples from participants in the National Health and Nutrition Examination Survey (NHANES) since 1999. More than 100 chemicals or their metabolites were measured in blood and urine samples from a subsample of participants from NHANES conducted by CDC's National Center for Health Statistics.

The National Human Exposure Assessment Survey (NHEXAS), sponsored by the U.S EPA in 1994, is a pilot and long-term study of the ways that humans are exposed to potentially toxic chemicals in the environment. The environmental chemicals include lead and arsenic, benzene and related volatile organic compounds, and various pesticides. The field of study covers the counties in Minnesota, Wisconsin, Illinois, Indiana, Ohio, and Michigan(47).

Some exposure data are available from investigation studies. For example, several studies tested urinary dimethylthiophosphate(DMTP) to evaluate organophosphorus(OP) pesticide exposure. These studies have provided information of OP pesticide exposure in apple thinners(48), children of agricultural pesticide applicators(49), children living in a large metropolitan area(50), living in an agricultural community(51), and living in farmworker households(52).

Because exposure data are limited, levels of exposure are often estimated through sophisticated modeling. The Hazardous Air Pollutant Exposure Model, Version 4 (HAPEM4) is such a type of model and has been used by the U. S. EPA to estimate inhalation exposure of air toxics for specified population groups. The National-Scale Air Toxics Assessment (NATA) database were developed with this exposure modeling and provides information about exposure concentration estimates of the 33 air pollutants at the census-tract, as well as county/state level. The goal of the national-scale assessment is to identify those air toxics which are of greatest potential concern, in terms of contribution to population risk. Currently, however, only two years of data are available, i.e. 1996 and 1999.

B.1.3 HEALTH OUTCOME DATA

Health outcomes of interest can include any outcomes along the timeline for the development natural history of a disease from preclinical measurements to death. These outcomes of interest are in the effort to link an environmental hazard or exposure with them in human beings. Sources of existing health outcome data include national, state and local level disease registries, vital statistics data, nation-wide health surveys, administrative data systems and large epidemiological cohort studies.

a). Centers for Disease Control and Prevention (CDC) Databases

The Centers for Disease Control and Prevention (CDC), as the lead federal agency for protecting the health and safety of people, has housed health data from birth and death records, medical records, interview surveys, through direct physical exams and laboratory testing. These databases maintained by the CDC could serve as valuable sources of health outcomes for the Environmental Public Health Tracking.

The National Vital Statistics System in the CDC's National Center of Health Statistics (NCHS) is one of the oldest and most successful databases. The database provides information about all vital events including births, deaths, marriages, divorces, and fetal deaths in the 50 States, 2 cities (Washington, DC, and New York City), and 5 United States territories (Puerto Rico, the Virgin Islands, Guam, American Samoa, and the Commonwealth of the Northern Mariana Islands).

The National Health and Nutrition Examination Survey (NHANES) is a major effort of the National Center for Health Statistics (NCHS) to assess the health and nutritional status of adults and children in the United States. The NHANES program began in the early 1960s and collects data on a nationally representative sample of about 5,000 persons each year. This survey consists of individual information on demographics, socioeconomic, dietary intake, and health. A physical examination conducts, which consists of medical and dental examination, physiological measurements, and laboratory tests administered by highly trained medical personnel.

The National Health Interview Survey (NHIS) is a multipurpose health survey conducted by the CDC's NCHS. The NHIS has been an important source of information about health and health care in the United States since it was first conducted in 1957.

The National Health Care Survey (NHCS) of the CDC's NCHS includes a series of health care provider surveys. These surveys provide information about the facilities that supply health care, the services rendered, and the characteristics of the patients served. This family of surveys includes National Ambulatory Medical Care Survey (NAMCS), National Hospital Ambulatory Medical Care Survey (NHAMCS), National Survey of Ambulatory Surgery (NSAS), National Hospital Discharge Survey (NHDS), National Nursing Home Survey (NNHS), National Home and Hospice Care Survey (NHHCS), National Employer Health Insurance Survey (NEHIS), and National Health Provider Inventory (NHPI). The detailed information of each survey refers to the CDC's NCHS website (<http://www.cdc.gov/nchs/nhcs.htm>).

The CDC's State and Local Area Integrated Telephone Survey (SLAITS) collects important health care data at State and local levels. It supplements current national data collection strategies by providing in-depth State and local area data to meet various program and policy needs in an ever-changing health care system. The SLAITS data included the Iowa and Washington State Health data (1997), Texas and Minnesota Child Well-Being and Welfare data (1997), National Survey of Early Childhood Health data (2000), National Survey of Children with Special Health Care Needs (2000-02, 2005-06), National Asthma Survey (2003), and National Survey of Children's Health (2003-2004).

The Longitudinal Studies of Aging (LSOAs) is a collaborative project between CDC's NCHS and the National Institute on Aging (NIA). The project consisted of four surveys: the 1984 Supplement on Aging, the 1984-1990 Longitudinal Study of Aging (LSOA), the 1994 Second Supplement on Aging (SOA II), and the 1994-2000 Second Longitudinal Study of Aging (LSOA II). The purpose of the project is to measure changes in the health, functional status,

living arrangements, and health services utilization among Americans of 70 years of age and over as they move into and through the oldest ages.

The CDC's National Program of Cancer Registries (NPCR) has been funding 45 states, District of Columbia and 3 U.S. territories to collect population-based cancer incidence data since 1994. The database provides information about patient demographics, tumor characteristics, state at diagnosis and first course of treatment. NPCR has four products for data release: U.S. Cancer Statistics (USCS), U.S. County Cancer Incidence Dataset, USCS Expanded Dataset, and USCS Restricted Access Datasets.

b). Other agency and institute databases

In addition, there are many databases of health outcomes maintained by other agencies and organizations. The following review provides a brief description of the most frequently assessed databases available from other agencies excluding CDC.

The Surveillance, Epidemiology, and End Results (SEER) Program of the National Cancer Institute (NCI) is a comprehensive source of population-based information on cancer incidence and survival in the United States. The program began in 1973 and collected data from population-based cancer registries covering approximately 26 percent of the US population. It provides information about stage of cancer at the time of diagnosis and patient survival data.

The Agency for Toxic Substances and Disease Registry (ATSDR) has maintained an active, state-based Hazard Substances Emergency Event Surveillance (HSEES) since 1990. This system provides the information of the public health consequences associated with the release of hazardous substances. It includes 15 participating states(53).

The American Association of Poison Control Centers (AAPCC) maintains the Toxic Exposure Surveillance System (TESS), which is the comprehensive poisoning surveillance

database and covered the entire population of the 50 states, the District of Columbia, and Puerto Rico in the United States. TESS is a valuable resource for environmental public health tracking. It provides information on the patient, the caller, the exposure, the substance(s), clinical toxicity, treatment, and medical outcome. Developed in 1983, TESS contains detailed toxicological information on 36.2 million cases reported to U.S. poison centers through 2003. That includes more than 2 million reports to poison centers for 2000 alone, an estimated 96% of all poison exposures reported to poison centers in the U.S (54, 55).

The National Cancer Database (NCDB) was established by the American Cancer Society (ACS) and the American College of Surgeons (ACoS) in 1989. The database served as a comprehensive clinical surveillance resource for cancer care in the United States. It captures 75% of all newly diagnosed cancer cases in the United States annually and contains over 15 million cases of reported cancer diagnoses for the period 1985 through 2002. The database provides information about patient characteristics, tumor staging and histology characteristics, type of first course treatment administered, disease recurrence, and survival information.

c). Large research cohort study databases

Large cohort studies are other potential valuable sources of existing data for environmental public health tracking. Data collected in cohort studies are at the individual level. Detailed individual information such as contact information and address, lifestyle information like smoking and diet, and other individual factors are usually available in the data of cohort studies. A brief introduction of some well-documented cohort studies will be given in the following section.

The Framingham study, directed by the National Heart, Lung, and Blood Institute (NHLBI), is designed to identify the common factors or characteristics that contribute to CVD among 5,209 participants, their offspring and third generation. The study lasted over 50 years(56).

The Nurse's Health Study, sponsored by the National Institutes of Health., is the largest prospective investigation of risk factors for major chronic diseases in women. The study included Nurses' Health Study and the Nurses' Health Study II. The Nurses' Health Study began in 1976 and approximately 122,000 nurses aged 30 to 55 from the 11 most populous states participated in the study(57). The Nurses' Health Study II began in 1989 and a total of 116,686 women between the ages of 25 and 42 years are involved in the study(58, 59).

The California Teachers Study, sponsored by the State of California, is a prospective study of 133,479 current and former public school teachers or administrators who participate in the California State Teachers Retirement System (STRS). The purpose of the study is to identify the risk factors associated with breast cancer research. The study began in the fall of 1995 (60).

The Agricultural Health Study was sponsored as a joint collaboration between the National Institutes of Health and the Environmental Protection Agency in 1994. Over 89,000 farmers as well as their families and commercial pesticide applicators from North Carolina and Iowa states participated in the project. The Study is designed to identify occupational, lifestyle, and genetic factors that may affect the rate of diseases in farming populations (61).

The American Cancer Society Cancer Prevention Study includes the Cancer Prevention Study I and the Cancer Prevention Study II. The Cancer Prevention Study I was a prospective mortality study and began in 1959. Over 1 millions participants were recruited into the study in twenty-five states and followed up through September 1972(62). The Cancer Prevention Study II (CPS II) includes the CPS II Baseline Cohort, CPS II Nutrition Cohort and CPS II Biospecimen

Repository. The Cancer Prevention Study II is to examine the impact of environmental and lifestyle factors on cancer etiology. Approximately 1.2 million men and women in 50 states, the District of Columbia, and Puerto Rico participated in the CPS-II Baseline Cohort in 1982. The data have been examined extensively in relation to cancer mortality(63, 64). The CPS-II Nutrition Cohort was established in 1992 and 1993 as a subgroup of the larger CPS-II Cohort. A total of 184,194 men and women in middle-aged to elderly living in 21 states in the United States participated in the cohort. The study is designed to evaluate the associations between dietary as well as other risk factors and cancer incidence (65). The CPS-II Biospecimen Cohort was initiated to obtain blood samples from 40,000 surviving members of the CPS-II Nutrition Cohort in 1998.

The Women's Health Initiative (WHI) Observational study was established by the National Institutes of Health (NIH) to address cardiovascular disease, cancer, and osteoporosis among women with age 50-79 years old. A total of 93,676 women were enrolled in this study (66).

The Cardiovascular Health Study (CHS) sponsored by the National Heart, Lung, and Blood Institute (NHLBI), is a population-based, longitudinal study of risk factors for the development and progression of CHD and stroke in the U.S. adults over the age of 65 years. The study started in 1989 and will end in 2009. There are 5,201 participants in this study (67).

The Alameda County Study began in 1995 and is designed to examine the risk factors associated with health and mortality among 6928 adults living in Alameda County, California. The last update of this study occurred on 1997(68).

B.1.4 SPATIAL DATA

Geospatial data has been increasingly used in public health research. Many agencies or organizations provide geospatial databases. Some of these databases are now readily available free of charge on the internet. In the following section, sources of spatial data available that are potentially useful for environmental public health tracking are briefly described.

The U.S. Geological Survey (USGS) has long been a source of geospatial data. Most of these data are now available in digital formats. The Digital Cartographic Data products provided by USGS include Digital Orthophoto Quadrangle (DOQ), Digital Raster Graphics (DRG), Digital Elevation Models (DEMs), Digital Line Graphs (DLGs) and National Hydrography Dataset (NHD). The Digital Line Graphs data set provides information on transportation lines, hydrography, political boundaries, and elevation contours, vegetable surface cover, manmade features, non-vegetative features as well as township, range, and section lines. The National Hydrography Dataset (NHD) is a comprehensive set of digital spatial data that contains information about surface water features such as lakes, ponds, streams, rivers, springs and wells.

The U.S. Census Bureau provides a number of spatial datasets. The Topologically integrated Geographic Encoding and Referencing (TIGER) system released by the U.S. Census Bureau contains information on street, railway, political boundaries of census, and major hydrographic features. It covers the 50 states, the District of Columbia, Puerto Rico, the Virgin Island and the outlying area of pacific area. The Cartographic Boundary Files are selected generalized extracts from the Census Bureau's TIGER geographic database and are designed for use in a Geographic Information System (GIS) or similar mapping system. The files are available in three formats: ARC/INFO EXPORT (.e00) format, ArcView Shapefile (.shp) format and ARC/INFO

Ungenerate (ASCII) format. The spatial resolutions in the file include census tract, block, division, region, zip code and metropolitan Areas etc.

The Environmental Systems Research Institute (ESRI) specializes in geographic information systems (GIS) software. The spatial datasets provided by ESRI include the StreetMap data, Community Data, Basemaps and census data. The StreetMap Pro and StreetMap Premium enhanced street datasets provide routing, geocoding, and high-quality cartographic display for the entire United States. The ESRI Community Data encompasses a variety of datasets that help companies and organizations analyze markets, profile customers, evaluate competitors, and more. ESRI software users can access most of these datasets at an additional cost.

B.2 STRATEGIES OF DATA LINKAGE

Data linkage is the task of accurately identifying records corresponding to the same entity from one or more data sources and the methodology for joining together corresponding records from two or more databases. The entity in public health studies usually refers to individual, family or geographical region etc. Using existing data for evaluating possible associations between environmental hazard or exposure and health outcome, linkages of these types of data are the intermediate and important step.

B.2.1 DATA LINKAGE

Linking health outcome with data of environmental hazards or exposures depends on following aspects: purpose of study, study design and nature and structure of available data. The

general steps involved in computer-based linkage procedures for two or more databases are shown as the following scheme (Fig B-2).

The accuracy of data linkage greatly depends on the linkage strategy, methodology and quality of data. The first step of data linkage is to be familiar with the databases and to determine data fields or variables in each database to be studied. Therefore, some nuisance data fields can be eliminated from databases. This process makes the databases to be linked more compact. The next step is to identify the key variables across the different databases. These key variables are used as linkage variables, which should exist in each database to be linked. The selections of key variables are vital to successful data linkage. The quality of these key variables such as missing values and typographic errors can heavily influence the accuracy and quality of linked dataset. Standardization of key variables is another vital process for data linkage. Without standardization, many truly matched records might result in unmatched records. Standardization includes converting common words into standard spelling, for example “road” or “street” in an address is converted into “Rd” or “St” for each database, using the same system of unit, and applying the same coding system for same variables etc.

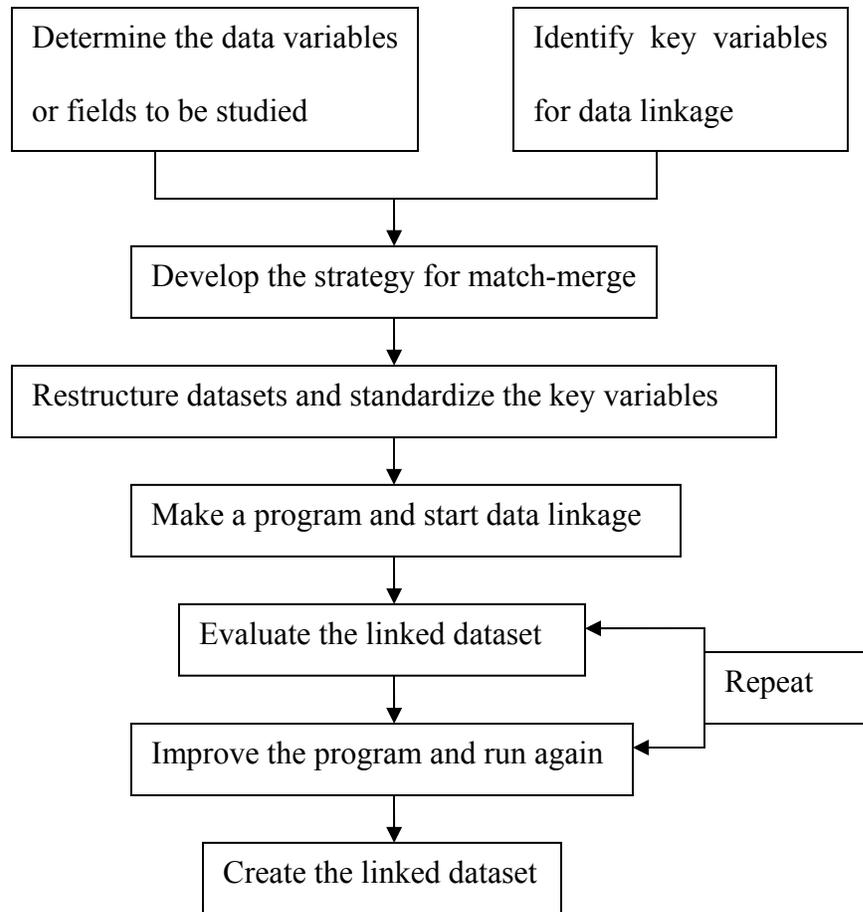


Figure B-2 The general procedure of data linkage

B.2.2 METHODS OF DATA LINKAGE

The linkage of routinely collected data is a relatively cheap and readily available method of obtaining data for monitoring population health and for research purpose. With increasing capacity of data storage and development in data linkage techniques, there has been substantial

increase in linking existing data to obtain information on large population in research or health surveillance. There are two main types of data linkage methods: deterministic data linkage and probabilistic data linkage(69).

a. Deterministic linkage

Deterministic linkage, also called exact linkage, is the ability to link records from different databases that exactly match on error-free identifying fields or variables. Such identifying fields should be precise, robust, stable over time and highly available. Deterministic linkage can be used if a unique identifier is available in the record fields of all the dataset to be linked. However, in individual level datasets, the identifiers to be used in deterministic linkage usually are confidential information, such as Social Security Number (SSN), name and date of birth in health dataset. It is clear that such linkage of two or more datasets can infringe on individual privacy. For datasets aggregated in geographic regions or time dimensions, the issue of a breach in confidentiality or privacy is less problematic. The key variables for these type datasets might usually be the variables of geographic region such as zip code, county and state or time such as day, month, and year. For example, in a time series study of air pollution, the linkage of monitored air pollution data with the number daily hospital admissions (counts) data on a given day is based on the key variable of calendar day and not any individual level data element (70, 71).

Several studies have been conducted to evaluate the performance of deterministic linkage on health data with different combinations of key words. Herman et al used deterministic linkage methods to link birth registry data with infant mortality and morbidity surveillance data sets from the United States (Georgia, Missouri, Utah and Washington State), Israel, Norway, Scotland and

Western Australia in 1997. A unique identifier, national identification numbers, were available for these datasets. The study reported that the methodology was highly successful (72).

Li et al completed data linkage with the population registry, hospital discharge and Vital Statistics registry datasets using deterministic linkage with a combination of key variables surname, sex and date of birth. This study found that the combination of surname, sex and date of birth appears to be optimal using deterministic linkage(73).

Simon et al merged records of young women with invasive breast cancer identified through three population-based cancer registries, to state birth certificate records with deterministic linkage. The performance of data linkage with different combinations of key variables was compared. The study found that SSN appears to be fairly accurate for linkage for linking cancer registries to other data sources (74).

The deterministic data linkage method is simple and easier to interpret. Deterministic data linkage can be accomplished with standard statistical software packages such as SPSS and SAS. However, deterministic linkage can be very time-consuming and very practitioner-dependent. In addition, it can not handle partial agreements easily, which may lead to false negative matching(75).

b. Probabilistic linkage

Data linkage is inevitably probabilistic in nature. Deterministic linkage is an extreme example of probabilistic linkage in that it is 100% certain that records of different databases, exactly matched on key variables, indeed refer to the same entity. However it may not true due to

duplication of identifiers and errors in creating or transmitting records. Thus deterministic linkage does not adequately reflect the uncertainty that may exist for some potential links. Moreover, unique and reliable identifiers may not be available in the databases to be linked in many situations. It is desirable that different linkage strategy should be developed to address these issues. Probabilistic linkage is the alternative to deterministic linkage and is developed when the rules of classic probability theory have been applied in data linkage.

The method of probabilistic linkage is defined as record linkage of two or more files that utilizes the probabilities of agreement and disagreement between a range of matching variables (69). Newcomber et al first reported the possibility of probabilistically linking records and provided crucial insights that led to computerized approaches to record linkage in the absence of unique identifiers(76). Subsequently, Fellegi and Sunter introduced the mathematical and statistical foundation to formalize the theory of probabilistic data linkage, building on the concept provided by Newcomber (77). The basic principle of probabilistic linkage is to use the probability of a true match ($m(\gamma)$) and the probability of a true non-match ($u(\gamma)$) based on matching variable to determine its (dis)agreement weight. The (dis)agreement weights for all matching variables were determined by the same strategy. The total weight for a given record link is simply the sum of the (dis)agreement weights for each matching variables. If all or most matching variables agree, the total weight will be a large positive number; otherwise, it is a small positive number. By comparing the total weight with threshold values, a true match, possible match or true non-match is determined (69, 75).

Probabilistic data linkage is commonly used in health research studies. Several studies have evaluated the accuracy of probabilistic linkage in matching records of health databases. Newgard et al evaluated the accuracy of using probabilistic linkage for matching de-identified ambulance

records to a state trauma registry. It found that probabilistic linkage is a valid method for matching ambulance records to a trauma registry without the use of patient identifiers but the accuracy is related to the selection of common variables included in the analysis (78).

Nitsch et al assessed the validation of probabilistic record linkage to identify births to a cohort of women by linking the females in the cohort to birth records held by the Scottish Maternity Record System (SMR 2) based on surname, maiden name, initials, date of birth and postcode. It concluded that probabilistic record linkage to routine maternity records applied to population-based cohort can have high specificity, and as such may be reliably used in epidemiological research(79).

A few studies have been conducted to compare deterministic and probabilistic method on their performance of data linkage. Tromp M et al used the Dutch National Midwife Registry, the Dutch National Obstetrics Registry and the Dutch National Pediatrics Registry to compare probabilistic and deterministic record linkage techniques. The study reported that Probabilistic linkage identified 80% more links than a full deterministic linkage approach. External validation revealed an error rate of less than 1% (80).

Roos and Wajda compared deterministic and probabilistic methods and found that probabilistic linkage had great advantages in those situations where only a moderate amount of extra information was available(81).

Therefore, many studies have applied the deterministic and probabilistic linkage strategies to combine health data across the world. After a comprehensive review of these studies, it might provide some clues of data linkage in using secondary data for environmental public health tracking.

B.3 STATISTICAL METHODS

The relationship between environmental exposure and health outcomes is complex. For example, multiple relationships between environmental exposures and health outcomes often exist, as multiple environmental factors may contribute to a single disease. On the other hand, a single environmental factor may contribute to many diseases (82). The type of statistical model that is used to evaluate this possible association requires a careful consideration. The selection of a statistical model is much dependent on the purpose of study, study design and the availability of dataset.

Mather et al organized the statistical models for the analysis of environmental and health data as well as the environment and health relationship into three groups. The first group is composed of descriptive analyses which describes trends, generates baselines and compare temporal and spatial changes of health and environmental data. The second group is an ecologic analysis which uses aggregated health and environmental data to provide information on the relationship between hazards and health outcomes. The third group consists of etiologic research to test the relationships between environmental factors and health outcomes (83). Following the statistical framework proposed by Mather, statistical methods are organized into two categories: descriptive analysis and etiologic analysis. A variety of available statistical methods which are appropriate for the analysis of environment and health data has been reviewed based on these two categories.

B.3.1 DESCRIPTIVE ANALYSIS

Descriptive analysis often represents the first step in describing or investigating a new event or condition. It can be used to monitor the trends in hazards, exposures and health outcomes and to search for clues of cause of disease.

Time trend: Trend analysis of environmental hazards is important for environmental public health tracking to characterize the background and changes in environmental contaminants. In addition, Kyle et al suggested that for cases where associations between environmental hazards and health outcomes have been well studied and established, it may be more relevant to focus on tracking environmental contaminants (82). In such cases, trend analysis of environmental hazards can provide information for planning and action of prevention of diseases. Examining time trend of health outcomes is helpful to understand the temporal variation of disease and to identify disease time clustering. Statistical approaches for detection of clustering over time include scan statistic which is the maximum number of observed cases in an interval of pre-selected length, as the interval is allowed to scan, or slide along, the time frame of interest (84), and Ederer-Myers-Mantel test in which the test statistic, m_1 will be large when cases are clustered; otherwise, m_1 will be small when cases occur uniformly through time(85, 86).

Disease cluster or clustering: Disease cluster or clustering is useful to highlight areas at apparently high risk and to provide useful etiologic clues to disease origins. Many methods have been proposed to detect disease cluster or clustering. The methods include nearest neighbor test or Cuzick-and-Edwards' test (87) that compare the distance between cases to expected distance(88), spatial autocorrelation which summarizes the degree to which similar observations tend to occur near each other(89) , and spatial scan statistic which identifies the collection of

diseases least consistent with the null hypothesis and provides a significance value representing the detected cluster's "unusualness" (90).

Visual analysis (mapping): Mapping is a technique to provide insight into spatial and spatial-temporal variation in exposure and disease risk. In public health, disease mapping has a long history. The earliest example is Snow's famous cholera map in the mid-19th century. Disease mapping has been widely applied in the field of health studies and was often used in cancer studies (91-93).

For a fair comparison, the rates from map are required to be adjusted for potential confounders such as age, sex and race. Standardization methods include direct standardization and indirect standardization. With indirect standardization, disease maps typically show standardized mortality or morbidity ratios (SMRs or SIRs). With direct standardization, the adjusted rate was used to map.

The disease rate based on a large population may be a better estimator. However, the rates from small population are likely to be elevated artificially, which reflect lack of data rather than true elevated risk. Spatial smoothing is one method for reducing the noise in rates associated with geographic region. Common approaches to spatial smoothing include locally weighted average in which a smoothed value is obtained by averaging the values associated with neighboring regions, empirical Bayes smoothing which use a Bayesian approach to define the analytic form of the compromise estimator(47), and kernel smoothing in which a kernel function and a defined bandwidth are used to controls the amount of smoothing (94)

Ecological analysis: In ecological analysis, the purpose is to evaluate temporal and/or geographic variations in exposure to environmental factors across population groups in relation to health outcomes measured on time and/or geographic scale. The approach is easily adapted to

use data which are routinely collected. The major limitation of the approach is ecologic fallacy, which is that an association between exposure and health outcome observed at aggregated group level may not validate at the individual level. For this reason, ecological analysis can be useful to generate but not confirm a hypothesis for the causality of disease.

Temporal ecological analysis has been applied to evaluate data aggregated in time scale such as daily, weekly, monthly or yearly. It has been widely used in air pollution studies. Time-series analysis of air pollution is this type of study, which assess the association between daily variation of air pollutants and daily count of health outcomes after controlling for climatic factors, seasonality and secular trend. The major two statistical methods of time-series analysis are generalized linear model (GLM) with parametric natural cubic splines (95, 96) and generalized additive models (GAM)(97, 98).

Geographic ecological analysis to evaluate the correlation between aggregate exposure and aggregate health outcome in geographical unit is usually defined by administrative boundary such as city, county and states. Traditional statistics such as linear, Poisson and logistic regression may not be appropriate for spatial data analysis without incorporating the information on neighborhood relationships and spatially correlated error terms. Spatial autoregressive models are developed by treating observations of the outcome variable at other locations as additional covariates in the model with associated parameters defining spatial association(99). The generalized linear mixed model is another method for modeling regional counts and rates developed by introducing random effects and hierarchical modeling into models (100). Bayesian models are developed to fit very complicated hierarchical models including those with spatially correlated random effects. The basic idea of this model is that of using the data to define a probability distribution for each of the model parameters, then using these distributions to draw

inference. Several studies report that Bayesian models offer dramatic improvements in statistical accuracy over conventional statistical methods(101, 102).

B.3.2 ETIOLOGICAL ANALYSIS

Etiological analysis usually uses exposure/hazard and health outcome data at the individual level. The most common study designs to evaluate the possible association between health outcomes and environmental hazards/exposures consist of a cohort study and a case-control study including a case-crossover study.

The strength of association between environmental factors and health outcome is not usually high. In addition, the association between environmental factors and health outcomes is complicate. The ability to detect an association is influenced by many factors. Lack of validity of testing can be expected if there are unidentified confounding factors, if confounding factors are measured with errors or if controlling for confounding factors is inadequate(103). Therefore, the use of appropriate statistical models is important to detect an association between environmental factor and health outcome.

In cohort study, survival analysis such as Cox proportion hazard model was often applied to estimate the relative risk. In a case-control study, the statistical method is logistic regression, which is used to estimate odds ratios. Conditional logistic regression is required to estimate odds ratios in paired case control study or case-crossover study, which we will discuss later (see proposed demonstrative studies).

B.4 ISSUES OF USING SECONDARY DATA FOR ENVIRONMENTAL PUBLIC HEALTH TRACKING

B.4.1 DATA ISSUES

Using secondary data for environmental public health tracking is promising and feasible. Development of the EPHT program depends on the availability, quality, timeliness, compatibility and utility of existing data. However, secondary data collected for specific aims may not be sufficient for environmental public health tracking. A few publications discussed about the strengths and limitations of using secondary data for epidemiologic studies (104-106). Sorensen et al provided a comprehensive review of issues of using secondary data sources for epidemiologic research, which include completeness of data, accuracy of information collected, registration period, data format, data accessibility and availability and possibilities of linkage with other data sources(107).

Completeness of data is concerned with the function of registry system and data quality. Cases can be missed even in the best registration system. Improving the registry system may enhance completeness of data. On the other hand, completeness of data is also related to the completeness of information on subjects collected in the system. Significant missing or incomplete data can negate the value of a secondary data source. Completeness of data can also refer to time dimension. For example, many EPA monitors operate for only short periods of time or have intermittent periods with no data collection. It yields gaps in the data. This compromises the value of the data.

Accuracy of information collected reflects data quality as well. Incorrect information in secondary data may come from incorrect data entry or lack of entry of available information and

errors of measurement, i.e. information which may not reflect the true condition or characteristic of an individual or condition. It is very difficult or impossible to evaluate the accuracy of information in secondary data.

Data from different time periods have some potential issues. The issues are related to something changing over time in the data collection process. For example, in the NEI database, a different method for estimating emissions will be applied. Data produced at different periods of time might not be comparable. For a disease registration system, disease definition, access to care, diagnostic technology, and disease reporting can change over time (108).

Data accessibility and availability and data format are the issues related to data sharing, data delivery and data application. If these issues are not well addressed, it might be very difficult or time and labor consuming to make use of secondary data.

B.4.2 ISSUES OF ACCURATELY IN ASSESSING ENVIRONMENTAL EXPOSURE

Accurately assessing environmental exposure is a major issue in environmental public health tracking. As mentioned before, available data systems lack adequate measures of human exposure. Many environmental agent exposures have not identified the specific and effective biologic indices. Assessment of human exposure will continue to be critical in environmental public health tracking. However, even if exposure data are available, these data usually provide information of intensity of exposure at single point of time, i.e. time of sample collection. To accurately assess environmental exposure, the data need to capture the duration, frequency and intensity of exposure over a specified time period. Moreover, timing of exposure is another important factor which could influence the correct estimation of exposure level. Rothman presented a useful structure to reveal the relation between timing exposure and the occurrence of disease. It

suggested that there would be a specific time period during which a given exposure is causally related to disease, i.e. time window of exposure. The exposure before or after the etiologically important time window might not be a good measurement of disease relevant exposure (109). Several publications reported the importance and methods for considering the time of period of exposure in addition to the dose of exposure in evaluating the effect of exposure on onset of diseases (110-112).

B.4.3 ISSUES OF CONFIDENTIALITY AND PRIVACY

Confidentiality and privacy are barriers to secondary data accessibility and availability. Health data usually include personal data information, which is important information for data linkage. Using these data might be an invasion of privacy and confidentiality. Moreover, it is difficult and potentially impossible to secure informed consent from individuals in identified in population-based administrative and registry data. Therefore, it is very important to clarify who has the right to use specific data and which authorities should approve the use of the data for environmental public health tracking. The strategies to address these issues should be developed in the future for balancing between the protection of privacy as well as confidentiality and data accessibility and availability.

In a word, all the above issues should be carefully considered when secondary data are used for environmental public health tracking.

B.5 REFERENCE:

1. Cantor KP, Hoover R, Mason TJ, McCabe LJ. Associations of cancer mortality with halomethanes in drinking water. *J Natl Cancer Inst* 1978;61:979-85.
2. Doyle TJ, Zheng W, Cerhan JR, et al. The association of drinking water source and chlorination by-products with cancer incidence among postmenopausal women in Iowa: a prospective cohort study. *Am J Public Health* 1997;87:1168-76.
3. Gottlieb MS, Carr JK, Morris DT. Cancer and drinking water in Louisiana: colon and rectum. *Int J Epidemiol* 1981;10:117-25.
4. Kanarek MS, Young TB. Drinking water treatment and risk of cancer death in Wisconsin. *Environ Health Perspect* 1982;46:179-86.
5. Koivusalo M, Hakulinen T, Vartiainen T, Pukkala E, Jaakkola JJ, Tuomisto J. Drinking water mutagenicity and urinary tract cancers: a population-based case-control study in Finland. *Am J Epidemiol* 1998;148:704-12.
6. Koivusalo M, Pukkala E, Vartiainen T, Jaakkola JJ, Hakulinen T. Drinking water chlorination and cancer-a historical cohort study in Finland. *Cancer Causes Control* 1997;8:192-200.
7. Morris RD, Audet AM, Angelillo IF, Chalmers TC, Mosteller F. Chlorination, chlorination by-products, and cancer: a meta-analysis. *Am J Public Health* 1992;82:955-63.
8. Pope CA, 3rd, Bates DV, Raizenne ME. Health effects of particulate air pollution: time for reassessment? *Environ Health Perspect* 1995;103:472-80.
9. Pope CA, 3rd, Burnett RT, Thun MJ, et al. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *Jama* 2002;287:1132-41.
10. Atkinson RW, Anderson HR, Sunyer J, et al. Acute effects of particulate air pollution on respiratory admissions: results from APHEA 2 project. *Air Pollution and Health: a European Approach*. *Am J Respir Crit Care Med* 2001;164:1860-6.
11. Daggett DA, Myers JD, Anderson HA. Ozone and particulate matter air pollution in Wisconsin: trends and estimates of health effects. *Wmj* 2000;99:47-51.
12. Schwartz J, Slater D, Larson TV, Pierson WE, Koenig JQ. Particulate air pollution and hospital emergency room visits for asthma in Seattle. *Am Rev Respir Dis* 1993;147:826-31.

13. Just J, Segala C, Sahraoui F, Priol G, Grimfeld A, Neukirch F. Short-term health effects of particulate and photochemical air pollution in asthmatic children. *Eur Respir J* 2002;20:899-906.
14. Whittemore AS, Korn EL. Asthma and air pollution in the Los Angeles area. *Am J Public Health* 1980;70:687-96.
15. Wang X, Ding H, Ryan L, Xu X. Association between air pollution and low birth weight: a community-based study. *Environ Health Perspect* 1997;105:514-20.
16. Hansen C, Neller A, Williams G, Simpson R. Low levels of ambient air pollution during pregnancy and fetal growth among term neonates in Brisbane, Australia. *Environ Res* 2006.
17. Rogers JF, Dunlop AL. Air pollution and very low birth weight infants: a target population? *Pediatrics* 2006;118:156-64.
18. Woodruff TJ, Grillo J, Schoendorf KC. The relationship between selected causes of postneonatal infant mortality and particulate air pollution in the United States. *Environ Health Perspect* 1997;105:608-12.
19. Patel AS, Talbott EO, Zborowski JV, et al. Risk of cancer as a result of community exposure to gasoline vapors. *Arch Environ Health* 2004;59:497-503.
20. Schnatter AR, Rosamilia K, Wojcik NC. Review of the literature on benzene exposure and leukemia subtypes. *Chem Biol Interact* 2005;153-154:9-21.
21. Sanderson KV. Arsenic and Skin Cancer. *Trans St Johns Hosp Dermatol Soc* 1963;49:115-22.
22. Wagner SL, Maliner JS, Morton WE, Braman RS. Skin cancer and arsenical intoxication from well water. *Arch Dermatol* 1979;115:1205-7.
23. el Khwsky F, Bedwani R, D'Avanzo B, et al. Risk factors for non-melanomatous skin cancer in Alexandria, Egypt. *Int J Cancer* 1994;56:375-8.
24. Karagas MR, Stukel TA, Morris JS, et al. Skin cancer risk in relation to toenail arsenic concentrations in a US population-based case-control study. *Am J Epidemiol* 2001;153:559-65.
25. Hopenhayn-Rich C, Biggs ML, Smith AH. Lung and kidney cancer mortality associated with arsenic in drinking water in Cordoba, Argentina. *Int J Epidemiol* 1998;27:561-9.
26. Hopenhayn-Rich C, Biggs ML, Fuchs A, et al. Bladder cancer mortality associated with arsenic in drinking water in Argentina. *Epidemiology* 1996;7:117-24.
27. Chiu HF, Ho SC, Wang LY, Wu TN, Yang CY. Does arsenic exposure increase the risk for liver cancer? *J Toxicol Environ Health A* 2004;67:1491-500.

28. Yang CY. Does arsenic exposure increase the risk of development of peripheral vascular diseases in humans? *J Toxicol Environ Health A* 2006;69:1797-804.
29. Hertz-Picciotto I, Arrighi HM, Hu SW. Does arsenic exposure increase the risk for circulatory disease? *Am J Epidemiol* 2000;151:174-81.
30. Chiu HF, Chang CC, Tsai SS, Yang CY. Does arsenic exposure increase the risk for diabetes mellitus? *J Occup Environ Med* 2006;48:63-7.
31. Hopenhayn C, Ferreccio C, Browning SR, et al. Arsenic exposure from drinking water and birth weight. *Epidemiology* 2003;14:593-602.
32. Ahmad SA, Sayed MH, Barua S, et al. Arsenic in drinking water and pregnancy outcomes. *Environ Health Perspect* 2001;109:629-31.
33. Telisman S, Prpic-Majic D, Beritic T. PbB and ALAD in mentally retarded and normal children. *Int Arch Occup Environ Health* 1983;52:361-9.
34. Shih RA, Glass TA, Bandeen-Roche K, et al. Environmental lead exposure and cognitive function in community-dwelling older adults. *Neurology* 2006.
35. Tellez-Rojo MM, Bellinger DC, Arroyo-Quiroz C, et al. Longitudinal associations between blood lead concentrations lower than 10 microg/dL and neurobehavioral development in environmentally exposed children in Mexico City. *Pediatrics* 2006;118:e323-30.
36. Medina-Ramon M, Zanobetti A, Cavanagh DP, Schwartz J. Extreme temperatures and mortality: assessing effect modification by personal characteristics and specific cause of death in a multi-city case-only analysis. *Environ Health Perspect* 2006;114:1331-6.
37. Conti S, Meli P, Minelli G, et al. Epidemiologic study of mortality during the Summer 2003 heat wave in Italy. *Environ Res* 2005;98:390-9.
38. Catelinois O, Rogel A, Laurier D, et al. Lung cancer attributable to indoor radon exposure in france: impact of the risk models and uncertainty analysis. *Environ Health Perspect* 2006;114:1361-6.
39. Field RW, Krewski D, Lubin JH, et al. An overview of the North American residential radon and lung cancer case-control studies. *J Toxicol Environ Health A* 2006;69:599-631.
40. Curcio G, Ferrara M, Moroni F, D'Inzeo G, Bertini M, De Gennaro L. Is the brain influenced by a phone call? An EEG study of resting wakefulness. *Neurosci Res* 2005;53:265-70.
41. Park SK, Ha M, Im HJ. Ecological study on residences in the vicinity of AM radio broadcasting towers and cancer death: preliminary observations in Korea. *Int Arch Occup Environ Health* 2004;77:387-94.

42. Ahlbom A, Day N, Feychting M, et al. A pooled analysis of magnetic fields and childhood leukaemia. *Br J Cancer* 2000;83:692-8.
43. Commission TPEh. *America's Environmental Health Gap: Why the County Needs a Nationwide Health Tracking Network*, 2000
44. CDC. *Environmental Public Health Tracking Program*. 2003.
45. Thacker SB, Stroup DF, Parrish RG, Anderson HA. Surveillance in environmental public health: issues, systems, and sources. *Am J Public Health* 1996;86:633-8.
46. McGeehin MA, Qualters JR, Niskar AS. National environmental public health tracking program: bridging the information gap. *Environ Health Perspect* 2004;112:1409-13.
47. Pellizzari E, Lioy P, Quackenboss J, et al. Population-based exposure measurements in EPA region 5: a phase I field study in support of the National Human Exposure Assessment Survey. *J Expo Anal Environ Epidemiol* 1995;5:327-58.
48. Simcox NJ, Camp J, Kalman D, et al. Farmworker exposure to organophosphorus pesticide residues during apple thinning in central Washington State. *Am Ind Hyg Assoc J* 1999;60:752-61.
49. Loewenherz C, Fenske RA, Simcox NJ, Bellamy G, Kalman D. Biological monitoring of organophosphorus pesticide exposure among children of agricultural workers in central Washington State. *Environ Health Perspect* 1997;105:1344-53.
50. Lu C, Knutson DE, Fisker-Andersen J, Fenske RA. Biological monitoring survey of organophosphorus pesticide exposure among pre-school children in the Seattle metropolitan area. *Environ Health Perspect* 2001;109:299-303.
51. Koch D, Lu C, Fisker-Andersen J, Jolley L, Fenske RA. Temporal association of children's pesticide exposure and agricultural spraying: report of a longitudinal biological monitoring study. *Environ Health Perspect* 2002;110:829-33.
52. Curl CL, Fenske RA, Kissel JC, et al. Evaluation of take-home organophosphorus pesticide exposure among agricultural workers and their children. *Environ Health Perspect* 2002;110:A787-92.
53. Orr MF, Ruckart PZ. Surveillance of hazardous substances releases due to system interruptions, 2002. *J Hazard Mater* 2006.
54. Watson WA, Litovitz TL, Belson MG, et al. The Toxic Exposure Surveillance System (TESS): Risk assessment and real-time toxicovigilance across United States poison centers. *Toxicol Appl Pharmacol* 2005;207:604-10.
55. Watson WA, Litovitz TL, Rodgers GC, Jr., et al. 2004 Annual report of the American Association of Poison Control Centers Toxic Exposure Surveillance System. *Am J Emerg Med* 2005;23:589-666.

56. Dawber TR, Meadors GF, Moore FE, Jr. Epidemiological approaches to heart disease: the Framingham Study. *Am J Public Health* 1951;41:279-81.
57. Belanger CF, Hennekens CH, Rosner B, Speizer FE. The nurses' health study. *Am J Nurs* 1978;78:1039-40.
58. Chasan-Taber L, Willett WC, Manson JE, et al. Prospective study of oral contraceptives and hypertension among women in the United States. *Circulation* 1996;94:483-9.
59. Grodstein F, Colditz GA, Hunter DJ, Manson JE, Willett WC, Stampfer MJ. A prospective study of symptomatic gallstones in women: relation with oral contraceptives and other risk factors. *Obstet Gynecol* 1994;84:207-14.
60. Horn-Ross PL, Hoggatt KJ, West DW, et al. Recent diet and breast cancer risk: the California Teachers Study (USA). *Cancer Causes Control* 2002;13:407-15.
61. Alavanja MC, Sandler DP, McMaster SB, et al. The Agricultural Health Study. *Environ Health Perspect* 1996;104:362-9.
62. Hammond EC. Cancer etiology: new prospective epidemiological study. *CA Cancer J Clin* 1959;9:177-8.
63. Underwood CR. Cancer prevention Study II. *J Med Assoc Ga* 1982;71:565-6.
64. Stellman SD, Garfinkel L. Smoking habits and tar levels in a new American Cancer Society prospective study of 1.2 million men and women. *J Natl Cancer Inst* 1986;76:1057-63.
65. Calle EE, Rodriguez C, Jacobs EJ, et al. The American Cancer Society Cancer Prevention Study II Nutrition Cohort: rationale, study design, and baseline characteristics. *Cancer* 2002;94:2490-501.
66. Design of the Women's Health Initiative clinical trial and observational study. The Women's Health Initiative Study Group. *Control Clin Trials* 1998;19:61-109.
67. Fried LP, Borhani NO, Enright P, et al. The Cardiovascular Health Study: design and rationale. *Ann Epidemiol* 1991;1:263-76.
68. Kotler P, Wingard DL. The effect of occupational, marital and parental roles on mortality: the Alameda County Study. *Am J Public Health* 1989;79:607-12.
69. Blakely T, Salmond C. Probabilistic record linkage and a method to calculate the positive predictive value. *Int J Epidemiol* 2002;31:1246-52.
70. Schwartz J. Daily deaths are associated with combustion particles rather than SO₂ in Philadelphia. *Occup Environ Med* 2000;57:692-7.

71. Schwartz J, Marcus A. Mortality and air pollution in London: a time series analysis. *Am J Epidemiol* 1990;131:185-94.
72. Herman AA, McCarthy BJ, Bakewell JM, et al. Data linkage methods used in maternally-linked birth and infant death surveillance data sets from the United States (Georgia, Missouri, Utah and Washington State), Israel, Norway, Scotland and Western Australia. *Paediatr Perinat Epidemiol* 1997;11 Suppl 1:5-22.
73. Li B, Quan H, Fong A, Lu M. Assessing record linkage between health care and Vital Statistics databases using deterministic methods. *BMC Health Serv Res* 2006;6:48.
74. Simon MS, Mueller BA, Deapen D, Copeland G. A comparison of record linkage yield for health research using different variable sets. *Breast Cancer Res Treat* 2005;89:107-10.
75. Gomatam S, Carter R, Ariet M, Mitchell G. An empirical comparison of record linkage procedures. *Stat Med* 2002;21:1485-96.
76. Newcombe HB, Kennedy JM, Axford SJ, James AP. Automatic linkage of vital records. *Science* 1959;130:954-9.
77. Fellegi I, Sunter A. A theory of record linkage. *J. AM. Statist. Assoc.* 1969;64:1183-1210.
78. Newgard CD. Validation of probabilistic linkage to match de-identified ambulance records to a state trauma registry. *Acad Emerg Med* 2006;13:69-75.
79. Nitsch D, Morton S, DeStavola BL, Clark H, Leon DA. How good is probabilistic record linkage to reconstruct reproductive histories? Results from the Aberdeen Children of the 1950s study. *BMC Med Res Methodol* 2006;6:15.
80. Tromp M, Meray N, Ravelli AC, Reitsma JB, Bonsel GJ. Medical Record Linkage of Anonymous Registries without Validated Sample Linkage of the Dutch Perinatal Registries. *Stud Health Technol Inform* 2005;116:125-30.
81. Roos LL, Wajda A. Record linkage strategies. Part I: Estimating information and evaluating approaches. *Methods Inf Med* 1991;30:117-23.
82. Kyle AD, Balmes JR, Buffler PA, Lee PR. Integrating research, surveillance, and practice in environmental public health tracking. *Environ Health Perspect* 2006;114:980-4.
83. Mather FJ, White LE, Langlois EC, et al. Statistical methods for linking health, exposure, and hazards. *Environ Health Perspect* 2004;112:1440-5.
84. Wallenstein S. A test for detection of clustering over time. *Am J Epidemiol* 1980;111:367-72.
85. Stark CR, Mantel N. Lack of seasonal- or temporal-spatial clustering of Down's syndrome births in Michigan. *Am J Epidemiol* 1967;86:199-213.

86. Ederer F, Myers E, Mantel N. A statistical problem in space and time: Do leukemia cases come in clusters? *Biometrics* 1964;20:623-626.
87. Cuzick J, Edwards R. Spatial clustering for inhomogeneous populations. *J. R. Stat. Soc., Ser. B, Methodol.* 1990;52:73-104.
88. Jacquez GM. A k nearest neighbour test for space-time interaction. *Stat Med* 1996;15:1935-49.
89. Ward MP, Carpenter TE. Techniques for analysis of disease clustering in space and in time in veterinary epidemiology. *Prev Vet Med* 2000;45:257-84.
90. Kulldorff M. A spatial scan statistic. *Communications in Statistics: Theory and Methods* 1997;26:1487-1496.
91. Christakos G, Lai JJ. A study of the breast cancer dynamics in North Carolina. *Soc Sci Med* 1997;45:1503-17.
92. Jarup L, Best N, Toledano MB, Wakefield J, Elliott P. Geographical epidemiology of prostate cancer in Great Britain. *Int J Cancer* 2002;97:695-9.
93. Manton KG, Woodbury MA, Stallard E, Riggan WB, Creason JP, Pellom AC. Empirical Bayes procedures for stabilizing maps of U.S. cancer mortality rates. *J Am Stat Assoc* 1989;84:637-50.
94. Lawson AB. Disease map reconstruction. *Stat Med* 2001;20:2183-204.
95. Chen Y, Yang Q, Krewski D, Shi Y, Burnett RT, McGrail K. Influence of relatively low level of particulate air pollution on hospitalization for COPD in elderly people. *Inhal Toxicol* 2004;16:21-5.
96. Dominici F, McDermott A, Daniels M, Zeger SL, Samet JM. Revised analyses of the National Morbidity, Mortality, and Air Pollution Study: mortality among residents of 90 cities. *J Toxicol Environ Health A* 2005;68:1071-92.
97. Arena VC, Mazumdar S, Zborowski JV, et al. A retrospective investigation of PM10 in ambient air and cardiopulmonary hospital admissions in Allegheny County, Pennsylvania: 1995-2000. *J Occup Environ Med* 2006;48:38-47.
98. Pitard A, Zeghnoun A, Courseaux A, et al. Short-term associations between air pollution and respiratory drug sales. *Environ Res* 2004;95:43-52.
99. Gelfand AE, Vounatsou P. Proper multivariate conditional autoregressive models for spatial data analysis. *Biostatistics* 2003;4:11-25.
100. Zhang H. On estimation and prediction for spatial generalized linear mixed models. *Biometrics* 2002;58:129-36.

101. Greenland S, Poole C. Empirical-Bayes and semi-Bayes approaches to occupational and environmental hazard surveillance. *Arch Environ Health* 1994;49:9-16.
102. Greenland S. Methods for epidemiologic analyses of multiple exposures: a review and comparative study of maximum-likelihood, preliminary-testing, and empirical-Bayes regression. *Stat Med* 1993;12:717-36.
103. Prentice RL, Thomas D. Methodologic research needs in environmental epidemiology: data analysis. *Environ Health Perspect* 1993;101 Suppl 4:39-48.
104. Bright RA, Avorn J, Everitt DE. Medicaid data as a resource for epidemiologic studies: strengths and limitations. *J Clin Epidemiol* 1989;42:937-45.
105. Lauderdale DS, Furner SE, Miles TP, Goldberg J. Epidemiologic uses of Medicare data. *Epidemiol Rev* 1993;15:319-27.
106. Roos LL, Mustard CA, Nicol JP, et al. Registries and administrative data: organization and accuracy. *Med Care* 1993;31:201-12.
107. Sorensen HT, Sabroe S, Olsen J. A framework for evaluation of secondary data sources for epidemiological research. *Int J Epidemiol* 1996;25:435-42.
108. Leibson CL, Ballard DJ, Whisnant JP, Melton LJ, 3rd. The compression of morbidity hypothesis: promise and pitfalls of using record-linked data bases to assess secular trends in morbidity and mortality. *Milbank Q* 1992;70:127-54.
109. Rothman KJ. Induction and latent periods. *Am J Epidemiol* 1981;114:253-9.
110. Thomas DC. Models for exposure-time-response relationships with applications to cancer epidemiology. *Annu Rev Public Health* 1988;9:451-82.
111. Checkoway H, Pearce N, Hickey JL, Dement JM. Latency analysis in occupational epidemiology. *Arch Environ Health* 1990;45:95-100.
112. White E, Hunt JR, Casso D. Exposure measurement in cohort studies: the challenges of prospective data collection. *Epidemiol Rev* 1998;20:43-56.

BIBLIOGRAPHY

- Adami J, Nyren O, Bergstrom R, Ekblom A, Engholm G, et al. 1998. Smoking and the risk of leukemia, lymphoma, and multiple myeloma (Sweden). *Cancer Causes Control* 9:49-56.
- Ahlbom A, Day N, Feychting M, Roman E, Skinner J, et al. 2000. A pooled analysis of magnetic fields and childhood leukaemia. *Br J Cancer* 83:692-8.
- Ahmad SA, Sayed MH, Barua S, Khan MH, Faruquee MH, et al. 2001. Arsenic in drinking water and pregnancy outcomes. *Environ Health Perspect* 109:629-31.
- Alavanja MC, Brownson RC, Lubin JH, Berger E, Chang J, Boice JD, Jr. 1994. Residential radon exposure and lung cancer among nonsmoking women. *J Natl Cancer Inst* 86:1829-37.
- Alavanja MC, Sandler DP, McMaster SB, Zahm SH, McDonnell CJ, et al. 1996. The Agricultural Health Study. *Environ Health Perspect* 104:362-9.
- Anderson HR, Spix C, Medina S, Schouten JP, Castellsague J, et al. 1997. Air pollution and daily admissions for chronic obstructive pulmonary disease in 6 European cities: results from the APHEA project. *Eur Respir J* 10:1064-71.
- Andrews JO, Tingen MS. 2006. The effect of smoking, smoking cessation, and passive smoke exposure on common laboratory values in clinical settings: a review of the evidence. *Crit Care Nurs Clin North Am* 18:63-9, xii.
- Arena VC, Mazumdar S, Zborowski JV, Talbott EO, He S, et al. 2006. A retrospective investigation of PM10 in ambient air and cardiopulmonary hospital admissions in Allegheny County, Pennsylvania: 1995-2000. *J Occup Environ Med* 48:38-47.
- Atkinson RW, Anderson HR, Sunyer J, Ayres J, Baccini M, et al. 2001. Acute effects of particulate air pollution on respiratory admissions: results from APHEA 2 project. *Air Pollution and Health: a European Approach. Am J Respir Crit Care Med* 164:1860-6.
- Austin H, Cole P. 1986. Cigarette smoking and leukemia. *J Chronic Dis* 39:417-21.
- Banks EC, Ferretti LE, Shucard DW. 1997. Effects of low level lead exposure on cognitive function in children: a review of behavioral, neuropsychological and biological evidence. *Neurotoxicology* 18:237-81.
- Barker DJ. 1995. The fetal and infant origins of disease. *Eur J Clin Invest* 25:457-63.
- Barker DJ. 2004. The developmental origins of adult disease. *J Am Coll Nutr* 23:588S-95S.

- Barnett AG, Williams GM, Schwartz J, Best TL, Neller AH, et al. 2006. The effects of air pollution on hospitalizations for cardiovascular disease in elderly people in Australian and new zealand cities. *Environ Health Perspect* 114:1018-23.
- Barnett AG, Williams GM, Schwartz J, Neller AH, Best TL, et al. 2005. Air pollution and child respiratory health: a case-crossover study in Australia and New Zealand. *Am J Respir Crit Care Med* 171:1272-8.
- Belanger CF, Hennekens CH, Rosner B, Speizer FE. 1978. The nurses' health study. *Am J Nurs* 78:1039-40.
- Benhamou S, Benhamou E, Tirmarche M, Flamant R. 1985. Lung cancer and use of cigarettes: a French case-control study. *J Natl Cancer Inst* 74:1169-75.
- Bhopal RS, Phillimore P, Moffatt S, Foy C. 1994. Is living near a coking works harmful to health? A study of industrial air pollution. *J Epidemiol Community Health* 48:237-47.
- Blakely T, Salmond C. 2002. Probabilistic record linkage and a method to calculate the positive predictive value. *Int J Epidemiol* 31:1246-52.
- Bobak M. 2000. Outdoor air pollution, low birth weight, and prematurity. *Environ Health Perspect* 108:173-6.
- Bobak M, Leon DA. 1999. Pregnancy outcomes and outdoor air pollution: an ecological study in districts of the Czech Republic 1986-8. *Occup Environ Med* 56:539-43.
- Bobak M, Richards M, Wadsworth M. 2001. Air pollution and birth weight in Britain in 1946. *Epidemiology* 12:358-9.
- Bosetti C, Gallus S, Garavello W, La Vecchia C. 2006. Smoking cessation and the risk of oesophageal cancer: An overview of published studies. *Oral Oncol* 42:957-64.
- Bright RA, Avorn J, Everitt DE. 1989. Medicaid data as a resource for epidemiologic studies: strengths and limitations. *J Clin Epidemiol* 42:937-45.
- Brook RD. 2005. You are what you breathe: evidence linking air pollution and blood pressure. *Curr Hypertens Rep* 7:427-34.
- Brown LM, Gibson R, Blair A, Burmeister LF, Schuman LM, et al. 1992. Smoking and risk of leukemia. *Am J Epidemiol* 135:763-8.
- Brownson RC. 1989. Cigarette smoking and risk of leukemia. *J Clin Epidemiol* 42:1025-8.
- Brownson RC, Chang JC, Davis JR. 1991. Cigarette smoking and risk of adult leukemia. *Am J Epidemiol* 134:938-41.

- Calle EE, Rodriguez C, Jacobs EJ, Almon ML, Chao A, et al. 2002. The American Cancer Society Cancer Prevention Study II Nutrition Cohort: rationale, study design, and baseline characteristics. *Cancer* 94:2490-501
- Cantor KP, Hoover R, Mason TJ, McCabe LJ. 1978. Associations of cancer mortality with halomethanes in drinking water. *J Natl Cancer Inst* 61:979-85.
- Catelinois O, Rogel A, Laurier D, Billon S, Hemon D, et al. 2006. Lung cancer attributable to indoor radon exposure in france: impact of the risk models and uncertainty analysis. *Environ Health Perspect* 114:1361-6.
- CDC. 2003. Environmental Public Health Tracking Program.
- Chang CC, Tsai SS, Ho SC, Yang CY. 2005. Air pollution and hospital admissions for cardiovascular disease in Taipei, Taiwan. *Environ Res* 98:114-9.
- Chang JS, Selvin S, Metayer C, Crouse V, Golembesky A, Buffler PA. 2006. Parental smoking and the risk of childhood leukemia. *Am J Epidemiol* 163:1091-100.
- Chasan-Taber L, Willett WC, Manson JE, Spiegelman D, Hunter DJ, et al. 1996. Prospective study of oral contraceptives and hypertension among women in the United States. *Circulation* 94:483-9.
- Checkoway H, Pearce N, Hickey JL, Dement JM. 1990. Latency analysis in occupational epidemiology. *Arch Environ Health* 45:95-100.
- Chelghoum Y, Danaila C, Belhabri A, Charrin C, Le QH, et al. 2002. Influence of cigarette smoking on the presentation and course of acute myeloid leukemia. *Ann Oncol* 13:1621-7.
- Chen L, Mengersen K, Tong S. 2007a. Spatiotemporal relationship between particle air pollution and respiratory emergency hospital admissions in Brisbane, Australia. *Sci Total Environ* 373:57-67.
- Chen L, Yang W, Jennison BL, Goodrich A, Omaye ST. 2002. Air pollution and birth weight in northern Nevada, 1991-1999. *Inhal Toxicol* 14:141-57.
- Chen TM, Gokhale J, Shofer S, Kuschner WG. 2007b. Outdoor Air Pollution: Particulate Matter Health Effects. *Am J Med Sci* 333:235-43.
- Chen Y, Yang Q, Krewski D, Shi Y, Burnett RT, McGrail K. 2004. Influence of relatively low level of particulate air pollution on hospitalization for COPD in elderly people. *Inhal Toxicol* 16:21-5.
- Chiu HF, Chang CC, Tsai SS, Yang CY. 2006. Does arsenic exposure increase the risk for diabetes mellitus? *J Occup Environ Med* 48:63-7.

- Chiu HF, Ho SC, Wang LY, Wu TN, Yang CY. 2004. Does arsenic exposure increase the risk for liver cancer? *J Toxicol Environ Health A* 67:1491-500.
- Christakos G, Lai JJ. 1997. A study of the breast cancer dynamics in North Carolina. *Soc Sci Med* 45:1503-17.
- Cohen AJ, Ross Anderson H, Ostro B, Pandey KD, Krzyzanowski M, et al. 2005. The global burden of disease due to outdoor air pollution. *J Toxicol Environ Health A* 68:1301-7.
- Commission TPEh. 2000 America's Environmental Health Gap: Why the County Needs a Nationwide Health Tracking Network.
- Conti S, Meli P, Minelli G, Solimini R, Toccaceli V, et al. 2005. Epidemiologic study of mortality during the Summer 2003 heat wave in Italy. *Environ Res* 98:390-9.
- Cook DG, Shaper AG, Pocock SJ, Kussick SJ. 1986. Giving up smoking and the risk of heart attacks. A report from The British Regional Heart Study. *Lancet* 2:1376-80.
- Curcio G, Ferrara M, Moroni F, D'Inzeo G, Bertini M, De Gennaro L. 2005. Is the brain influenced by a phone call? An EEG study of resting wakefulness. *Neurosci Res* 53:265-70.
- Curl CL, Fenske RA, Kissel JC, Shirai JH, Moate TF, et al. 2002. Evaluation of take-home organophosphorus pesticide exposure among agricultural workers and their children. *Environ Health Perspect* 110:A787-92.
- Cuzick J, Edwards R. 1990. Spatial clustering for inhomogeneous populations. *J. R. Stat. Soc., Ser. B, Methodol.* 52:73-104.
- D'Ippoliti D, Forastiere F, Ancona C, Agabiti N, Fusco D, et al. 2003. Air pollution and myocardial infarction in Rome: a case-crossover analysis. *Epidemiology* 14:528-35.
- Daggett DA, Myers JD, Anderson HA. 2000. Ozone and particulate matter air pollution in Wisconsin: trends and estimates of health effects. *Wmj* 99:47-51.
- Davis RM, Novotny TE. 1989. The epidemiology of cigarette smoking and its impact on chronic obstructive pulmonary disease. *Am Rev Respir Dis* 140:S82-4.
- Dawber TR, Meadors GF, Moore FE, Jr. 1951. Epidemiological approaches to heart disease: the Framingham Study. *Am J Public Health* 41:279-81.
- Dejmek J, Selevan SG, Benes I, Solansky I, Sram RJ. 1999. Fetal growth and maternal exposure to particulate matter during pregnancy. *Environ Health Perspect* 107:475-80.
- Dejmek J, Solansky I, Benes I, Lenicek J, Sram RJ. 2000. The impact of polycyclic aromatic hydrocarbons and fine particles on pregnancy outcome. *Environ Health Perspect* 108:1159-64.

- Delfino RJ, Sioutas C, Malik S. 2005. Potential role of ultrafine particles in associations between airborne particle mass and cardiovascular health. *Environ Health Perspect* 113:934-46.
- Design of the Women's Health Initiative clinical trial and observational study. The Women's Health Initiative Study Group. 1998. *Control Clin Trials* 19:61-109.
- Dockery DW, Pope CA, 3rd, Xu X, Spengler JD, Ware JH, et al. 1993. An association between air pollution and mortality in six U.S. cities. *N Engl J Med* 329:1753-9.
- Dolk H, Thakrar B, Walls P, Landon M, Grundy C, et al. 1999. Mortality among residents near cokeworks in Great Britain. *Occup Environ Med* 56:34-40.
- Dominici F. 2004. Time-series analysis of air pollution and mortality: a statistical review. *Res Rep Health Eff Inst*:3-27; discussion 9-33.
- Dominici F, McDermott A, Daniels M, Zeger SL, Samet JM. 2005. Revised analyses of the National Morbidity, Mortality, and Air Pollution Study: mortality among residents of 90 cities. *J Toxicol Environ Health A* 68:1071-92.
- Dominici F, McDermott A, Zeger SL, Samet JM. 2002. On the use of generalized additive models in time-series studies of air pollution and health. *Am J Epidemiol* 156:193-203.
- Donaldson K, Stone V, Seaton A, MacNee W. 2001. Ambient particle inhalation and the cardiovascular system: potential mechanisms. *Environ Health Perspect* 109 Suppl 4:523-7.
- Doyle TJ, Zheng W, Cerhan JR, Hong CP, Sellers TA, et al. 1997. The association of drinking water source and chlorination by-products with cancer incidence among postmenopausal women in Iowa: a prospective cohort study. *Am J Public Health* 87:1168-76.
- Drinker P. 1953. Deaths during the severe fog in London and environs, Dec. 5 to 9, 1952. *A M A Arch Ind Hyg Occup Med* 7:UNKNOWN.
- Dugandzic R, Dodds L, Stieb D, Smith-Doiron M. 2006. The association between low level exposures to ambient air pollution and term low birth weight: a retrospective cohort study. *Environ Health* 5:3.
- Ederer F, Myers E, Mantel N. 1964. A statistical problem in space and time: Do leukemia cases come in clusters? *Biometrics* 20:623-6.
- el Khwsky F, Bedwani R, D'Avanzo B, Assaad S, el Shafei Ali A, et al. 1994. Risk factors for non-melanomatous skin cancer in Alexandria, Egypt. *Int J Cancer* 56:375-8.
- Engeland A, Andersen A, Haldorsen T, Tretli S. 1996. Smoking habits and risk of cancers other than lung cancer: 28 years' follow-up of 26,000 Norwegian men and women. *Cancer Causes Control* 7:497-506.
- Fellegi I, Sunter A. 1969. A theory of record linkage. *J. AM. Statist. Assoc.* 64:1183-210.

- Ferreccio C, Sancha AM. 2006. Arsenic exposure and its impact on health in Chile. *J Health Popul Nutr* 24:164-75.
- Field RW, Krewski D, Lubin JH, Zielinski JM, Alavanja M, et al. 2006. An overview of the North American residential radon and lung cancer case-control studies. *J Toxicol Environ Health A* 69:599-631.
- Fried LP, Borhani NO, Enright P, Furberg CD, Gardin JM, et al. 1991. The Cardiovascular Health Study: design and rationale. *Ann Epidemiol* 1:263-76.
- Friedman GD. 1993. Cigarette smoking, leukemia, and multiple myeloma. *Ann Epidemiol* 3:425-8.
- Fubini B, Fenoglio I, Ceschino R, Ghiazza M, Martra G, et al. 2004. Relationship between the state of the surface of four commercial quartz flours and their biological activity in vitro and in vivo. *Int J Hyg Environ Health* 207:89-104.
- Gadgil AS, Jadhav RS. 2004. Street-level concentrations of suspended particulate matter (SPM), nitrogen dioxide (NO₂) and sulphur dioxide (SO₂) in Pune City. *J Environ Sci Eng* 46:143-50.
- Gehring U, Heinrich J, Kramer U, Grote V, Hochadel M, et al. 2006. Long-term exposure to ambient air pollution and cardiopulmonary mortality in women. *Epidemiology* 17:545-51.
- Gelfand AE, Vounatsou P. 2003. Proper multivariate conditional autoregressive models for spatial data analysis. *Biostatistics* 4:11-25.
- Gilboa SM, Mendola P, Olshan AF, Langlois PH, Savitz DA, et al. 2005. Relation between ambient air quality and selected birth defects, seven county study, Texas, 1997-2000. *Am J Epidemiol* 162:238-52.
- Glorennec P, Monroux F. 2007. Health impact assessment of PM₁₀ exposure in the city of Caen, France. *J Toxicol Environ Health A* 70:359-64.
- Gomatam S, Carter R, Ariet M, Mitchell G. 2002. An empirical comparison of record linkage procedures. *Stat Med* 21:1485-96.
- Gottlieb MS, Carr JK, Morris DT. 1981. Cancer and drinking water in Louisiana: colon and rectum. *Int J Epidemiol* 10:117-25.
- Gouveia N, Bremner SA, Novaes HM. 2004. Association between ambient air pollution and birth weight in Sao Paulo, Brazil. *J Epidemiol Community Health* 58:11-7.
- Greenburg L, Erhardt C, Field F, Reed JI, Seriff NS. 1963. Intermittent Air Pollution Episode in New York City, 1962. *Public Health Rep* 78:1061-4.

- Greenland S. 1993. Methods for epidemiologic analyses of multiple exposures: a review and comparative study of maximum-likelihood, preliminary-testing, and empirical-Bayes regression. *Stat Med* 12:717-36.
- Greenland S, Poole C. 1994. Empirical-Bayes and semi-Bayes approaches to occupational and environmental hazard surveillance. *Arch Environ Health* 49:9-16.
- Grodstein F, Colditz GA, Hunter DJ, Manson JE, Willett WC, Stampfer MJ. 1994. A prospective study of symptomatic gallstones in women: relation with oral contraceptives and other risk factors. *Obstet Gynecol* 84:207-14.
- Ha EH, Hong YC, Lee BE, Woo BH, Schwartz J, Christiani DC. 2001. Is air pollution a risk factor for low birth weight in Seoul? *Epidemiology* 12:643-8.
- Hales CN. 1997. Fetal and infant origins of adult disease. *J Clin Pathol* 50:359.
- Hammond EC. 1959. Cancer etiology: new prospective epidemiological study. *CA Cancer J Clin* 9:177-8.
- Hansen C, Neller A, Williams G, Simpson R. 2006. Low levels of ambient air pollution during pregnancy and fetal growth among term neonates in Brisbane, Australia. *Environ Res* .
- Health effects of outdoor air pollution. Committee of the Environmental and Occupational Health Assembly of the American Thoracic Society. 1996. *Am J Respir Crit Care Med* 153:3-50.
- Helfand WH, Lazarus J, Theerman P. 2001. Donora, Pennsylvania: an environmental disaster of the 20th century. *Am J Public Health* 91:553.
- Herman AA, McCarthy BJ, Bakewell JM, Ward RH, Mueller BA, et al. 1997. Data linkage methods used in maternally-linked birth and infant death surveillance data sets from the United States (Georgia, Missouri, Utah and Washington State), Israel, Norway, Scotland and Western Australia. *Paediatr Perinat Epidemiol* 11 Suppl 1:5-22.
- Hertz-Picciotto I. 1996. Comment: Toward a coordinated system for the surveillance of environmental health hazards. *Am J Public Health* 86:638-41.
- Hertz-Picciotto I, Arrighi HM, Hu SW. 2000. Does arsenic exposure increase the risk for circulatory disease? *Am J Epidemiol* 151:174-81.
- Hinwood AL, De Klerk N, Rodriguez C, Jacoby P, Runnion T, et al. 2006. The relationship between changes in daily air pollution and hospitalizations in Perth, Australia 1992-1998: a case-crossover study. *Int J Environ Health Res* 16:27-46.
- Holland WW, Bennett AE, Cameron IR, Florey CV, Leeder SR, et al. 1979. Health effects of particulate pollution: reappraising the evidence. *Am J Epidemiol* 110:527-659.

- Hopenhayn-Rich C, Biggs ML, Fuchs A, Bergoglio R, Tello EE, et al. 1996. Bladder cancer mortality associated with arsenic in drinking water in Argentina. *Epidemiology* 7:117-24.
- Hopenhayn-Rich C, Biggs ML, Smith AH. 1998. Lung and kidney cancer mortality associated with arsenic in drinking water in Cordoba, Argentina. *Int J Epidemiol* 27:561-9.
- Hopenhayn C, Ferreccio C, Browning SR, Huang B, Peralta C, et al. 2003. Arsenic exposure from drinking water and birth weight. *Epidemiology* 14:593-602.
- Horn-Ross PL, Hoggatt KJ, West DW, Krone MR, Stewart SL, et al. 2002. Recent diet and breast cancer risk: the California Teachers Study (USA). *Cancer Causes Control* 13:407-15.
- Jacquez GM. 1996. A k nearest neighbour test for space-time interaction. *Stat Med* 15:1935-49.
- Jarup L, Best N, Toledano MB, Wakefield J, Elliott P. 2002. Geographical epidemiology of prostate cancer in Great Britain. *Int J Cancer* 97:695-9.
- Jedrychowski W, Bendkowska I, Flak E, Penar A, Jacek R, et al. 2004. Estimated risk for altered fetal growth resulting from exposure to fine particles during pregnancy: an epidemiologic prospective cohort study in Poland. *Environ Health Perspect* 112:1398-402.
- Jurewicz J, Hanke W. 2006. Exposure to pesticides and childhood cancer risk: has there been any progress in epidemiological studies? *Int J Occup Med Environ Health* 19:152-69.
- Just J, Segala C, Sahraoui F, Priol G, Grimfeld A, Neukirch F. 2002. Short-term health effects of particulate and photochemical air pollution in asthmatic children. *Eur Respir J* 20:899-906.
- Kabat GC, Augustine A, Hebert JR. 1988. Smoking and adult leukemia: a case-control study. *J Clin Epidemiol* 41:907-14.
- Kalucka S. 2006. The occurrence of chronic obstructive pulmonary disease (COPD) in cigarette smoking families. *Przegl Lek* 63:848-57.
- Kan H, Chen B. 2003. A case-crossover analysis of air pollution and daily mortality in Shanghai. *J Occup Health* 45:119-24.
- Kan H, London SJ, Chen G, Zhang Y, Song G, et al. 2007. Differentiating the effects of fine and coarse particles on daily mortality in Shanghai, China. *Environ Int* 33:376-84.
- Kanarek MS, Young TB. 1982. Drinking water treatment and risk of cancer death in Wisconsin. *Environ Health Perspect* 46:179-86.
- Karagas MR, Stukel TA, Morris JS, Tosteson TD, Weiss JE, et al. 2001. Skin cancer risk in relation to toenail arsenic concentrations in a US population-based case-control study. *Am J Epidemiol* 153:559-65.

- Kasim K, Levallois P, Abdous B, Auger P, Johnson KC. 2005. Lifestyle factors and the risk of adult leukemia in Canada. *Cancer Causes Control* 16:489-500.
- Katsouyanni K, Touloumi G, Spix C, Schwartz J, Balducci F, et al. 1997. Short-term effects of ambient sulphur dioxide and particulate matter on mortality in 12 European cities: results from time series data from the APHEA project. *Air Pollution and Health: a European Approach. Bmj* 314:1658-63.
- Kestner JF, Jr. 1985. Cigarette smoking and lung cancer: a modern tragedy. *Del Med J* 57:443-5.
- Kettunen J, Lanki T, Tiittanen P, Aalto PP, Koskentalo T, et al. 2007. Associations of fine and ultrafine particulate air pollution with stroke mortality in an area of low air pollution levels. *Stroke* 38:918-22.
- Knaapen AM, Borm PJ, Albrecht C, Schins RP. 2004. Inhaled particles and lung cancer. Part A: Mechanisms. *Int J Cancer* 109:799-809.
- Koch D, Lu C, Fisker-Andersen J, Jolley L, Fenske RA. 2002. Temporal association of children's pesticide exposure and agricultural spraying: report of a longitudinal biological monitoring study. *Environ Health Perspect* 110:829-33.
- Koivusalo M, Hakulinen T, Vartiainen T, Pukkala E, Jaakkola JJ, Tuomisto J. 1998. Drinking water mutagenicity and urinary tract cancers: a population-based case-control study in Finland. *Am J Epidemiol* 148:704-12.
- Koivusalo M, Pukkala E, Vartiainen T, Jaakkola JJ, Hakulinen T. 1997. Drinking water chlorination and cancer-a historical cohort study in Finland. *Cancer Causes Control* 8:192-200.
- Kotler P, Wingard DL. 1989. The effect of occupational, marital and parental roles on mortality: the Alameda County Study. *Am J Public Health* 79:607-12.
- Krall EA, Dietrich T, Nunn ME, Garcia RI. 2006. Risk of tooth loss after cigarette smoking cessation. *Prev Chronic Dis* 3:A115.
- Krewski D, Burnett R, Jerrett M, Pope CA, Rainham D, et al. 2005. Mortality and long-term exposure to ambient air pollution: ongoing analyses based on the American Cancer Society cohort. *J Toxicol Environ Health A* 68:1093-109.
- Kulldorff M. 1997. A spatial scan statistic. *Communications in Statistics: Theory and Methods* 26:1487-96.
- Kunzli N, Tager IB. 2000. Long-term health effects of particulate and other ambient air pollution: research can progress faster if we want it to. *Environ Health Perspect* 108:915-8.
- Kunzli N, Tager IB. 2005. Air pollution: from lung to heart. *Swiss Med Wkly* 135:697-702.

- Kwon HJ, Cho SH, Nyberg F, Pershagen G. 2001. Effects of ambient air pollution on daily mortality in a cohort of patients with congestive heart failure. *Epidemiology* 12:413-9.
- Kyle AD, Balmes JR, Buffler PA, Lee PR. 2006. Integrating research, surveillance, and practice in environmental public health tracking. *Environ Health Perspect* 114:980-4.
- La Vecchia C. 1985. Patterns of cigarette smoking and trends in lung cancer mortality in Italy. *J Epidemiol Community Health* 39:157-64.
- Lapidus L, Bengtsson C, Lindquist O, Sigurdsson JA, Rafnsson V. 1986. Smoking--a risk factor for cardiovascular disease in women? Considerations based on a prospective population study of women in Gothenburg, Sweden. *Scand J Prim Health Care* 4:219-24.
- Lauderdale DS, Furner SE, Miles TP, Goldberg J. 1993. Epidemiologic uses of Medicare data. *Epidemiol Rev* 15:319-27.
- Lawson AB. 2001. Disease map reconstruction. *Stat Med* 20:2183-204.
- Lee BE, Ha EH, Park HS, Kim YJ, Hong YC, et al. 2003. Exposure to air pollution during different gestational phases contributes to risks of low birth weight. *Hum Reprod* 18:638-43.
- Lee JT, Kim H, Schwartz J. 2000. Bidirectional case-crossover studies of air pollution: bias from skewed and incomplete waves. *Environ Health Perspect* 108:1107-11
- Lee JT, Schwartz J. 1999. Reanalysis of the effects of air pollution on daily mortality in Seoul, Korea: A case-crossover design. *Environ Health Perspect* 107:633-6.
- Lee PN, Fry JS, Forey BA. 1990. Trends in lung cancer, chronic obstructive lung disease, and emphysema death rates for England and Wales 1941-85 and their relation to trends in cigarette smoking. *Thorax* 45:657-65.
- Leibson CL, Ballard DJ, Whisnant JP, Melton LJ, 3rd. 1992. The compression of morbidity hypothesis: promise and pitfalls of using record-linked data bases to assess secular trends in morbidity and mortality. *Milbank Q* 70:127-54.
- Levy D, Sheppard L, Checkoway H, Kaufman J, Lumley T, et al. 2001. A case-crossover analysis of particulate matter air pollution and out-of-hospital primary cardiac arrest. *Epidemiology* 12:193-9.
- Li B, Quan H, Fong A, Lu M. 2006. Assessing record linkage between health care and Vital Statistics databases using deterministic methods. *BMC Health Serv Res* 6:48.
- Li N, Sioutas C, Cho A, Schmitz D, Misra C, et al. 2003. Ultrafine particulate pollutants induce oxidative stress and mitochondrial damage. *Environ Health Perspect* 111:455-60.

- Liberti L, Notarnicola M, Primerano R, Zannetti P. 2006. Air pollution from a large steel factory: polycyclic aromatic hydrocarbon emissions from coke-oven batteries. *J Air Waste Manag Assoc* 56:255-60.
- Lichtenstein P, Holm NV, Verkasalo PK, Iliadou A, Kaprio J, et al. 2000. Environmental and heritable factors in the causation of cancer--analyses of cohorts of twins from Sweden, Denmark, and Finland. *N Engl J Med* 343:78-85.
- Lin CC, Santolaya-Forgas J. 1998. Current concepts of fetal growth restriction: part I. Causes, classification, and pathophysiology. *Obstet Gynecol* 92:1044-55.
- Lin M, Stieb DM, Chen Y. 2005. Coarse particulate matter and hospitalization for respiratory infections in children younger than 15 years in Toronto: a case-crossover analysis. *Pediatrics* 116:e235-40.
- Liu S, Krewski D, Shi Y, Chen Y, Burnett RT. 2003. Association between gaseous ambient air pollutants and adverse pregnancy outcomes in Vancouver, Canada. *Environ Health Perspect* 111:1773-8.
- Loewenherz C, Fenske RA, Simcox NJ, Bellamy G, Kalman D. 1997. Biological monitoring of organophosphorus pesticide exposure among children of agricultural workers in central Washington State. *Environ Health Perspect* 105:1344-53.
- Logan WP. 1953. Mortality in the London fog incident, 1952. *Lancet* 1:336-8.
- Logan WP. 1956. Mortality from fog in London, January, 1956. *Br Med J* 1:722-5.
- Lu C, Knutson DE, Fisker-Andersen J, Fenske RA. 2001. Biological monitoring survey of organophosphorus pesticide exposure among pre-school children in the Seattle metropolitan area. *Environ Health Perspect* 109:299-303.
- Mabuchi K, Bross DS, Kessler, II. 1985. Cigarette smoking and nasopharyngeal carcinoma. *Cancer* 55:2874-6.
- Maclure M. 1991. The case-crossover design: a method for studying transient effects on the risk of acute events. *Am J Epidemiol* 133:144-53.
- Maisonet M, Bush TJ, Correa A, Jaakkola JJ. 2001. Relation between ambient air pollution and low birth weight in the Northeastern United States. *Environ Health Perspect* 109 Suppl 3:351-6.
- Maisonet M, Correa A, Misra D, Jaakkola JJ. 2004. A review of the literature on the effects of ambient air pollution on fetal growth. *Environ Res* 95:106-15.
- Mannes T, Jalaludin B, Morgan G, Lincoln D, Sheppard V, Corbett S. 2005. Impact of ambient air pollution on birth weight in Sydney, Australia. *Occup Environ Med* 62:524-30.

- Manton KG, Woodbury MA, Stallard E, Riggan WB, Creason JP, Pellom AC. 1989. Empirical Bayes procedures for stabilizing maps of U.S. cancer mortality rates. *J Am Stat Assoc* 84:637-50.
- Marozienne L, Grazuleviciene R. 2002. Maternal exposure to low-level air pollution and pregnancy outcomes: a population-based study. *Environ Health* 1:6.
- Marrett LD, Meigs JW, Flannery JT. 1985. Bladder cancer in Connecticut: the role of cigarette smoking and other risk factors. *Conn Med* 49:718-26.
- Mather FJ, White LE, Langlois EC, Shorter CF, Swalm CM, et al. 2004. Statistical methods for linking health, exposure, and hazards. *Environ Health Perspect* 112:1440-5.
- McGeehin MA, Qualters JR, Niskar AS. 2004. National environmental public health tracking program: bridging the information gap. *Environ Health Perspect* 112:1409-13.
- McLaughlin JK, Hrubec Z, Linet MS, Heineman EF, Blot WJ, Fraumeni JF, Jr. 1989. Cigarette smoking and leukemia. *J Natl Cancer Inst* 81:1262-3.
- Medina-Ramon M, Zanobetti A, Cavanagh DP, Schwartz J. 2006a. Extreme temperatures and mortality: assessing effect modification by personal characteristics and specific cause of death in a multi-city case-only analysis. *Environ Health Perspect* 114:1331-6.
- Medina-Ramon M, Zanobetti A, Schwartz J. 2006b. The effect of ozone and PM10 on hospital admissions for pneumonia and chronic obstructive pulmonary disease: a national multicity study. *Am J Epidemiol* 163:579-88.
- Milei J, Mann J, Storino R, Molina HA. 1986. Relationship of cigarette smoking to acute myocardial infarction. A clinical-pathological study. *Arq Bras Cardiol* 47:247-52.
- Mills PK, Newell GR, Beeson WL, Fraser GE, Phillips RL. 1990. History of cigarette smoking and risk of leukemia and myeloma: results from the Adventist health study. *J Natl Cancer Inst* 82:1832-6.
- Monarca S, Donato F, Zerbini I, Calderon RL, Craun GF. 2006. Review of epidemiological studies on drinking water hardness and cardiovascular diseases. *Eur J Cardiovasc Prev Rehabil* 13:495-506.
- Mongelli M, Biswas A. 2001. A fetal growth standard derived from multiple modalities. *Early Hum Dev* 60:171-7.
- Moore A. 2005. Changing patterns of childhood mortality in Wolverhampton. *Arch Dis Child* 90:687-91.
- Morley R. 2006. Fetal origins of adult disease. *Semin Fetal Neonatal Med* 11:73-8.
- Morris RD, Audet AM, Angelillo IF, Chalmers TC, Mosteller F. 1992. Chlorination, chlorination by-products, and cancer: a meta-analysis. *Am J Public Health* 82:955-63.

- Mucci LA, Bjorkman L, Douglass CW, Pedersen NL. 2005. Environmental and heritable factors in the etiology of oral diseases--a population-based study of Swedish twins. *J Dent Res* 84:800-5.
- Navarro Silvera SA, Rohan TE. 2007. Trace elements and cancer risk: a review of the epidemiologic evidence. *Cancer Causes Control* 18:7-27.
- Neas LM, Schwartz J, Dockery D. 1999. A case-crossover analysis of air pollution and mortality in Philadelphia. *Environ Health Perspect* 107:629-31.
- Nemery B, Hoet PH, Nemmar A. 2001. The Meuse Valley fog of 1930: an air pollution disaster. *Lancet* 357:704-8.
- Neuberger JS, Gesell TF. 2002. Residential radon exposure and lung cancer: risk in nonsmokers. *Health Phys* 83:1-18.
- Newcombe HB, Kennedy JM, Axford SJ, James AP. 1959. Automatic linkage of vital records. *Science* 130:954-9.
- Newgard CD. 2006. Validation of probabilistic linkage to match de-identified ambulance records to a state trauma registry. *Acad Emerg Med* 13:69-75.
- Nicol CJ, Harrison ML, Laposa RR, Gimelshtein IL, Wells PG. 1995. A teratologic suppressor role for p53 in benzo[a]pyrene-treated transgenic p53-deficient mice. *Nat Genet* 10:181-7.
- Nitsch D, Morton S, DeStavola BL, Clark H, Leon DA. 2006. How good is probabilistic record linkage to reconstruct reproductive histories? Results from the Aberdeen Children of the 1950s study. *BMC Med Res Methodol* 6:15.
- Nyberg F, Pershagen G. 2000. Epidemiologic studies on the health effects of ambient particulate air pollution. *Scand J Work Environ Health* 26 Suppl 1:49-89.
- O'Brien JM. 2000. Environmental and heritable factors in the causation of cancer: analyses of cohorts of twins from Sweden, Denmark, and Finland, by P. Lichtenstein, N.V. Holm, P.K. Verkasalo, A. Iliadou, J. Kaprio, M. Koskenvuo, E. Pukkala, A. Skytthe, and K. Hemminki. *N Engl J Med* 343:78-84, 2000. *Surv Ophthalmol* 45:167-8.
- Olden K, Guthrie J. 2001. Genomics: implications for toxicology. *Mutat Res* 473:3-10.
- Orr MF, Ruckart PZ. 2006. Surveillance of hazardous substances releases due to system interruptions, 2002. *J Hazard Mater*.
- Osmond C, Barker DJ. 2000. Fetal, infant, and childhood growth are predictors of coronary heart disease, diabetes, and hypertension in adult men and women. *Environ Health Perspect* 108 Suppl 3:545-53.

- Paffenbarger RS, Jr., Brand RJ, Sholtz RI, Jung DL. 1978. Energy expenditure, cigarette smoking, and blood pressure level as related to death from specific diseases. *Am J Epidemiol* 108:12-8.
- Park SK, Ha M, Im HJ. 2004. Ecological study on residences in the vicinity of AM radio broadcasting towers and cancer death: preliminary observations in Korea. *Int Arch Occup Environ Health* 77:387-94.
- Parker JD, Woodruff TJ, Basu R, Schoendorf KC. 2005. Air pollution and birth weight among term infants in California. *Pediatrics* 115:121-8.
- Pasqualetti P, Festuccia V, Acitelli P, Collacciani A, Giusti A, Casale R. 1997. Tobacco smoking and risk of haematological malignancies in adults: a case-control study. *Br J Haematol* 97:659-62.
- Patel AS, Talbott EO, Zborowski JV, Rycheck JA, Dell D, et al. 2004. Risk of cancer as a result of community exposure to gasoline vapors. *Arch Environ Health* 59:497-503.
- Pellizzari E, Liroy P, Quackenboss J, Whitmore R, Clayton A, et al. 1995. Population-based exposure measurements in EPA region 5: a phase I field study in support of the National Human Exposure Assessment Survey. *J Expo Anal Environ Epidemiol* 5:327-58.
- Perera FP, Rauh V, Whyatt RM, Tang D, Tsai WY, et al. 2005. A summary of recent findings on birth outcomes and developmental effects of prenatal ETS, PAH, and pesticide exposures. *Neurotoxicology* 26:573-87.
- Pershagen G, Liang ZH, Hrubec Z, Svensson C, Boice JD, Jr. 1992. Residential radon exposure and lung cancer in Swedish women. *Health Phys* 63:179-86.
- Peters A. 2005. Particulate matter and heart disease: Evidence from epidemiological studies. *Toxicol Appl Pharmacol* 207:477-82.
- Peters A, Dockery DW, Muller JE, Mittleman MA. 2001. Increased particulate air pollution and the triggering of myocardial infarction. *Circulation* 103:2810-5.
- Peters A, Perz S, Doring A, Stieber J, Koenig W, Wichmann HE. 1999. Increases in heart rate during an air pollution episode. *Am J Epidemiol* 150:1094-8.
- Peters A, Pope CA, 3rd. 2002. Cardiopulmonary mortality and air pollution. *Lancet* 360:1184-5.
- Pitard A, Zeghnoun A, Courseaux A, Lamberty J, Delmas V, et al. 2004. Short-term associations between air pollution and respiratory drug sales. *Environ Res* 95:43-52.
- Pogoda JM, Preston-Martin S, Nichols PW, Ross RK. 2002. Smoking and risk of acute myeloid leukemia: results from a Los Angeles County case-control study. *Am J Epidemiol* 155:546-53.

- Pope CA, 3rd. 1989. Respiratory disease associated with community air pollution and a steel mill, Utah Valley. *Am J Public Health* 79:623-8.
- Pope CA, 3rd. 2000. Epidemiology of fine particulate air pollution and human health: biologic mechanisms and who's at risk? *Environ Health Perspect* 108 Suppl 4:713-23.
- Pope CA, 3rd, Bates DV, Raizenne ME. 1995. Health effects of particulate air pollution: time for reassessment? *Environ Health Perspect* 103:472-80.
- Pope CA, 3rd, Burnett RT, Thun MJ, Calle EE, Krewski D, et al. 2002. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *Jama* 287:1132-41.
- Pope CA, 3rd, Burnett RT, Thurston GD, Thun MJ, Calle EE, et al. 2004a. Cardiovascular mortality and long-term exposure to particulate air pollution: epidemiological evidence of general pathophysiological pathways of disease. *Circulation* 109:71-7.
- Pope CA, 3rd, Hansen ML, Long RW, Nielsen KR, Eatough NL, et al. 2004b. Ambient particulate air pollution, heart rate variability, and blood markers of inflammation in a panel of elderly subjects. *Environ Health Perspect* 112:339-45.
- Prentice RL, Thomas D. 1993. Methodologic research needs in environmental epidemiology: data analysis. *Environ Health Perspect* 101 Suppl 4:39-48.
- Rebelakos A, Trichopoulos D, Tzonou A, Zavitsanos X, Velonakis E, Trichopoulos A. 1985. Tobacco smoking, coffee drinking, and occupation as risk factors for bladder cancer in Greece. *J Natl Cancer Inst* 75:455-61.
- Reynolds P, Urayama KY, Von Behren J, Feusner J. 2004. Birth characteristics and hepatoblastoma risk in young children. *Cancer* 100:1070-6.
- Rice DC. 1996. Behavioral effects of lead: commonalities between experimental and epidemiologic data. *Environ Health Perspect* 104 Suppl 2:337-51.
- Risom L, Moller P, Loft S. 2005. Oxidative stress-induced DNA damage by particulate air pollution. *Mutat Res* 592:119-37.
- Ritz B, Yu F. 1999. The effect of ambient carbon monoxide on low birth weight among children born in southern California between 1989 and 1993. *Environ Health Perspect* 107:17-25.
- Ritz B, Yu F, Fruin S, Chapa G, Shaw GM, Harris JA. 2002. Ambient air pollution and risk of birth defects in Southern California. *Am J Epidemiol* 155:17-25.
- Rogers JF, Dunlop AL. 2006. Air pollution and very low birth weight infants: a target population? *Pediatrics* 118:156-64.

- Rogers JF, Thompson SJ, Addy CL, McKeown RE, Cowen DJ, Decoufle P. 2000. Association of very low birth weight with exposures to environmental sulfur dioxide and total suspended particulates. *Am J Epidemiol* 151:602-13.
- Roos LL, Mustard CA, Nicol JP, McLerran DF, Malenka DJ, et al. 1993. Registries and administrative data: organization and accuracy. *Med Care* 31:201-12.
- Roos LL, Wajda A. 1991. Record linkage strategies. Part I: Estimating information and evaluating approaches. *Methods Inf Med* 30:117-23.
- Rosenlund M, Berglind N, Pershagen G, Hallqvist J, Jonson T, Bellander T. 2006. Long-term exposure to urban air pollution and myocardial infarction. *Epidemiology* 17:383-90.
- Rothman KJ. 1981. Induction and latent periods. *Am J Epidemiol* 114:253-9.
- Ruidavets JB, Cournot M, Cassadou S, Giroux M, Meybeck M, Ferrieres J. 2005. Ozone air pollution is associated with acute myocardial infarction. *Circulation* 111:563-9.
- Saldiva PH, Pope CA, 3rd, Schwartz J, Dockery DW, Lichtenfels AJ, et al. 1995. Air pollution and mortality in elderly people: a time-series study in Sao Paulo, Brazil. *Arch Environ Health* 50:159-63.
- Sanderson KV. 1963. Arsenic and Skin Cancer. *Trans St Johns Hosp Dermatol Soc* 49:115-22.
- Sandler DP, Shore DL, Anderson JR, Davey FR, Arthur D, et al. 1993. Cigarette smoking and risk of acute leukemia: associations with morphology and cytogenetic abnormalities in bone marrow. *J Natl Cancer Inst* 85:1994-2003.
- Schins RP, Lightbody JH, Borm PJ, Shi T, Donaldson K, Stone V. 2004. Inflammatory effects of coarse and fine particulate matter in relation to chemical and biological constituents. *Toxicol Appl Pharmacol* 195:1-11.
- Schnatter AR, Rosamilia K, Wojcik NC. 2005. Review of the literature on benzene exposure and leukemia subtypes. *Chem Biol Interact* 153-154:9-21.
- Schwartz J. 2000. Daily deaths are associated with combustion particles rather than SO₂ in Philadelphia. *Occup Environ Med* 57:692-7.
- Schwartz J, Marcus A. 1990. Mortality and air pollution in London: a time series analysis. *Am J Epidemiol* 131:185-94.
- Schwartz J, Slater D, Larson TV, Pierson WE, Koenig JQ. 1993. Particulate air pollution and hospital emergency room visits for asthma in Seattle. *Am Rev Respir Dis* 147:826-31.
- Seaton A, Soutar A, Crawford V, Elton R, McNerlan S, et al. 1999. Particulate air pollution and the blood. *Thorax* 54:1027-32.

- Severson RK, Buckley JD, Woods WG, Benjamin D, Robison LL. 1993. Cigarette smoking and alcohol consumption by parents of children with acute myeloid leukemia: an analysis within morphological subgroups--a report from the Childrens Cancer Group. *Cancer Epidemiol Biomarkers Prev* 2:433-9.
- Severson RK, Davis S, Heuser L, Daling JR, Thomas DB. 1990. Cigarette smoking and acute nonlymphocytic leukemia. *Am J Epidemiol* 132:418-22.
- Shi L, Chia SE. 2001. A review of studies on maternal occupational exposures and birth defects, and the limitations associated with these studies. *Occup Med (Lond)* 51:230-44.
- Shih RA, Glass TA, Bandeen-Roche K, Carlson MC, Bolla KI, et al. 2006. Environmental lead exposure and cognitive function in community-dwelling older adults. *Neurology*.
- Simcox NJ, Camp J, Kalman D, Stebbins A, Bellamy G, et al. 1999. Farmworker exposure to organophosphorus pesticide residues during apple thinning in central Washington State. *Am Ind Hyg Assoc J* 60:752-61.
- Simon MS, Mueller BA, Deapen D, Copeland G. 2005. A comparison of record linkage yield for health research using different variable sets. *Breast Cancer Res Treat* 89:107-10.
- Sorensen HT, Sabroe S, Olsen J. 1996. A framework for evaluation of secondary data sources for epidemiological research. *Int J Epidemiol* 25:435-42.
- Speer SA, Semenza JC, Kurosaki T, Anton-Culver H. 2002. Risk factors for acute myeloid leukemia and multiple myeloma: a combination of GIS and case-control studies. *J Environ Health* 64:9-16; quiz 35-6.
- Spitz MR, Fueger JJ, Newell GR, Keating MJ. 1990. Leukemia and cigarette smoking. *Cancer Causes Control* 1:195-6.
- Spokas K, Graff C, Morcet M, Aran C. 2003. Implications of the spatial variability of landfill emission rates on geospatial analyses. *Waste Manag* 23:599-607.
- Sram RJ, Binkova B, Dejmek J, Bobak M. 2005. Ambient air pollution and pregnancy outcomes: a review of the literature. *Environ Health Perspect* 113:375-82.
- Stark CR, Mantel N. 1967. Lack of seasonal- or temporal-spatial clustering of Down's syndrome births in Michigan. *Am J Epidemiol* 86:199-213.
- Stellman SD, Garfinkel L. 1986. Smoking habits and tar levels in a new American Cancer Society prospective study of 1.2 million men and women. *J Natl Cancer Inst* 76:1057-63.
- Sullivan J, Sheppard L, Schreuder A, Ishikawa N, Siscovick D, Kaufman J. 2005. Relation between short-term fine-particulate matter exposure and onset of myocardial infarction. *Epidemiology* 16:41-8.

- Sunyer J, Basagana X. 2001. Particles, and not gases, are associated with the risk of death in patients with chronic obstructive pulmonary disease. *Int J Epidemiol* 30:1138-40.
- Sunyer J, Basagana X, Belmonte J, Anto JM. 2002. Effect of nitrogen dioxide and ozone on the risk of dying in patients with severe asthma. *Thorax* 57:687-93.
- Sunyer J, Schwartz J, Tobias A, Macfarlane D, Garcia J, Anto JM. 2000. Patients with chronic obstructive pulmonary disease are at increased risk of death associated with urban particle air pollution: a case-crossover analysis. *Am J Epidemiol* 151:50-6.
- Talbott EO, Youk AO, McHugh-Pemu KP, Zborowski JV. 2003. Long-term follow-up of the residents of the Three Mile Island accident area: 1979-1998. *Environ Health Perspect* 111:341-8.
- Talbott EO, Youk AO, McHugh KP, Shire JD, Zhang A, et al. 2000. Mortality among the residents of the Three Mile Island accident area: 1979-1992. *Environ Health Perspect* 108:545-52.
- Tapio S, Grosche B. 2006. Arsenic in the aetiology of cancer. *Mutat Res* 612:215-46.
- Telisman S, Prpic-Majic D, Beritic T. 1983. PbB and ALAD in mentally retarded and normal children. *Int Arch Occup Environ Health* 52:361-9.
- Tellez-Rojo MM, Bellinger DC, Arroyo-Quiroz C, Lamadrid-Figueroa H, Mercado-Garcia A, et al. 2006. Longitudinal associations between blood lead concentrations lower than 10 microg/dL and neurobehavioral development in environmentally exposed children in Mexico City. *Pediatrics* 118:e323-30.
- Thacker SB, Stroup DF, Parrish RG, Anderson HA. 1996. Surveillance in environmental public health: issues, systems, and sources. *Am J Public Health* 86:633-8.
- Thomas DC. 1988. Models for exposure-time-response relationships with applications to cancer epidemiology. *Annu Rev Public Health* 9:451-82.
- Tromp M, Meray N, Ravelli AC, Reitsma JB, Bonsel GJ. 2005. Medical Record Linkage of Anonymous Registries without Validated Sample Linkage of the Dutch Perinatal Registries. *Stud Health Technol Inform* 116:125-30.
- Tsai SS, Cheng MH, Chiu HF, Wu TN, Yang CY. 2006. Air pollution and hospital admissions for asthma in a tropical city: Kaohsiung, Taiwan. *Inhal Toxicol* 18:549-54.
- Tsai SS, Goggins WB, Chiu HF, Yang CY. 2003a. Evidence for an association between air pollution and daily stroke admissions in Kaohsiung, Taiwan. *Stroke* 34:2612-6.
- Tsai SS, Huang CH, Goggins WB, Wu TN, Yang CY. 2003b. Relationship between air pollution and daily mortality in a tropical city: Kaohsiung, Taiwan. *J Toxicol Environ Health A* 66:1341-9.

- Tsai SS, Yu HS, Liu CC, Yang CY. 2003c. Increased incidence of preterm delivery in mothers residing in an industrialized area in Taiwan. *J Toxicol Environ Health A* 66:987-94.
- Underwood CR. 1982. Cancer prevention Study II. *J Med Assoc Ga* 71:565-6.
- Ury HK. 1968. Photochemical air pollution and automobile accidents in Los Angeles. An investigation of oxidant and accidents, 1963 and 1965. *Arch Environ Health* 17:334-42.
- Wagner SL, Maliner JS, Morton WE, Braman RS. 1979. Skin cancer and arsenical intoxication from well water. *Arch Dermatol* 115:1205-7.
- Wallace L. 1996. Environmental exposure to benzene: an update. *Environ Health Perspect* 104 Suppl 6:1129-36.
- Wallenstein S. 1980. A test for detection of clustering over time. *Am J Epidemiol* 111:367-72.
- Wang J, Zhu L, Shen X. 2003. [PAHs pollution in air of coke plant and health risk assessment]. *Huan Jing Ke Xue* 24:135-8.
- Wang X, Ding H, Ryan L, Xu X. 1997. Association between air pollution and low birth weight: a community-based study. *Environ Health Perspect* 105:514-20.
- Ward MP, Carpenter TE. 2000. Techniques for analysis of disease clustering in space and in time in veterinary epidemiology. *Prev Vet Med* 45:257-84.
- Watson WA, Litovitz TL, Belson MG, Funk Wolkin AB, Patel M, et al. 2005a. The Toxic Exposure Surveillance System (TESS): Risk assessment and real-time toxicovigilance across United States poison centers. *Toxicol Appl Pharmacol* 207:604-10.
- Watson WA, Litovitz TL, Rodgers GC, Jr., Klein-Schwartz W, Reid N, et al. 2005b. 2004 Annual report of the American Association of Poison Control Centers Toxic Exposure Surveillance System. *Am J Emerg Med* 23:589-666.
- Weisberg E. 1985. Smoking and reproductive health. *Clin Reprod Fertil* 3:175-86.
- Wellenius GA, Schwartz J, Mittleman MA. 2005. Air pollution and hospital admissions for ischemic and hemorrhagic stroke among medicare beneficiaries. *Stroke* 36:2549-53.
- Wellenius GA, Schwartz J, Mittleman MA. 2006. Particulate air pollution and hospital admissions for congestive heart failure in seven United States cities. *Am J Cardiol* 97:404-8.
- White E, Hunt JR, Casso D. 1998. Exposure measurement in cohort studies: the challenges of prospective data collection. *Epidemiol Rev* 20:43-56.
- Whittemore AS, Korn EL. 1980. Asthma and air pollution in the Los Angeles area. *Am J Public Health* 70:687-96.

- Wietlisbach V, Pope CA, 3rd, Ackermann-Liebrich U. 1996. Air pollution and daily mortality in three Swiss urban areas. *Soz Praventivmed* 41:107-15.
- Wilhelm M, Ritz B. 2005. Local variations in CO and particulate air pollution and adverse birth outcomes in Los Angeles County, California, USA. *Environ Health Perspect* 113:1212-21.
- Woodruff TJ, Grillo J, Schoendorf KC. 1997. The relationship between selected causes of postneonatal infant mortality and particulate air pollution in the United States. *Environ Health Perspect* 105:608-12.
- Xie Y, Davies SM, Xiang Y, Robison LL, Ross JA. 2003. Trends in leukemia incidence and survival in the United States (1973-1998). *Cancer* 97:2229-35.
- Xu X, Ding H, Wang X. 1995. Acute effects of total suspended particles and sulfur dioxides on preterm delivery: a community-based cohort study. *Arch Environ Health* 50:407-15.
- Yang CY. 2006. Does arsenic exposure increase the risk of development of peripheral vascular diseases in humans? *J Toxicol Environ Health A* 69:1797-804.
- Yang CY, Chang CC, Chuang HY, Tsai SS, Wu TN, Ho CK. 2004. Relationship between air pollution and daily mortality in a subtropical city: Taipei, Taiwan. *Environ Int* 30:519-23.
- Yang CY, Tseng YT, Chang CC. 2003. Effects of air pollution on birth weight among children born between 1995 and 1997 in Kaohsiung, Taiwan. *J Toxicol Environ Health A* 66:807-16.
- Yorifuji T, Yamamoto E, Tsuda T, Kawakami N. 2005. Health impact assessment of particulate matter in Tokyo, Japan. *Arch Environ Occup Health* 60:179-85.
- Zanobetti A, Schwartz J. 2005. The effect of particulate air pollution on emergency admissions for myocardial infarction: a multicity case-crossover analysis. *Environ Health Perspect* 113:978-82.
- Zeger SL, Thomas D, Dominici F, Samet JM, Schwartz J, et al. 2000. Exposure measurement error in time-series studies of air pollution: concepts and consequences. *Environ Health Perspect* 108:419-26.
- Zhang H. 2002. On estimation and prediction for spatial generalized linear mixed models. *Biometrics* 58:129-36.
- Zhang M, Song Y, Cai X. 2007. A health-based assessment of particulate air pollution in urban areas of Beijing in 2000-2004. *Sci Total Environ* 376:100-8.
- Zimmerman D, Pavlik A, Ruggles A, Armstrong M. 1999. An experimental comparison of ordinary and universal kriging and inverse distance. *Mathematical Geology* 31:375-89.