Journal of Exposure Science and Environmental Epidemiology (2009) 19, 45–58

www.nature.com/jes

Use of health information in air pollution health research: Past successes and emerging needs

GEORGE D. THURSTON[a], MARNI Y.V. BEKKEDAL[b], ERIC M. ROBERTS[c], KAZUHIKO ITO[a], C. ARDEN POPE III[d], BARBARA S. GLENN[e], HALÛK ÖZKAYNAK[f] AND MARK J. UTELL[g]

[a]*Department of Environmental Medicine, New York University School of Medicine, Tuxedo Park, New York, USA*
[b]*Wisconsin Department of Health and Family Services, Madison, Wisconsin, USA*
[c]*Public Health Institute, Oakland, California, USA*
[d]*Department of Economics, Brigham Young University, Provo, Utah, USA*
[e]*National Center for Environmental Research, Office of Research and Development, US Environmental Protection Agency, Washington DC, USA*
[f]*National Exposure Research Laboratory, Office of Research and Development, US Environmental Protection Agency, RTP, North Carolina, USA*
[g]*Pulmonary Division, Department of Medicine and Environmental Medicine, University of Rochester Medical Center, Rochester, New York, USA*

In September 2006, the US Environmental Protection Agency and the US Centers for Disease Control (CDC) co-organized a symposium on "Air Pollution Exposure and Health." The main objective of this symposium was to identify opportunities for improving the use of exposure and health information in future studies of air pollution health effects. This paper deals with the health information needs of such studies. We begin with a selected review of different types of health data and how they were used in previous epidemiologic studies of health effects of ambient particulate matter (PM). We then examine the current and emerging information needs of the environmental health community, dealing with PM and other air pollutants of health concern. We conclude that the past use of routinely collected health data proved to be essential for activities to protect public health, including the identification and evaluation of health hazards by air pollution research, setting standards for criteria pollutants, surveillance of health outcomes to identify incidence trends, and the more recent CDC environmental public health tracking program. Unfortunately, access to vital statistics records that have informed such pivotal research has recently been curtailed sharply, threatening the continuation of the type of research necessary to support future standard setting and research on emerging exposure and health problems (e.g. asthma, multiple sclerosis, diabetes, and others), as well as our ability to evaluate the efficacy of regulatory and other prevention activities. A comprehensive devoted effort, perhaps new legislation, will be needed to address the standardization, centralization, and sharing of data sets, as well as to harmonize the interpretation of confidentiality and privacy protections across jurisdictions. These actions, combined with assuring researchers and public health practitioners appropriate access to data for evaluation of environmental risks, will be essential for the achievement of our environmental health protection goals.
Journal of Exposure Science and Environmental Epidemiology (2009) **19,** 45–58; doi:10.1038/jes.2008.41; published online 10 September 2008

Keywords: *air pollution, tracking, health data sets, epidemiology, accountability, particulate matter.*

Introduction

The US Environmental Protection Agency (EPA) and the US Centers for Disease Control (CDC) co-organized a symposium on "Air Pollution Exposure and Health" at RTP, NC, USA on 19–20 September 2006. The main goal of this symposium was to facilitate an interdisciplinary dialog among the air quality, air exposure, and health scientists for the purposes of improving exposure and health data sets and their linkage, for future air pollution epidemiology research, accountability research and environmental public health tracking (EPHT). Many significant advances in air pollution epidemiology have been made by studies that used administratively collected data on health outcomes. The linkage of air pollution levels with cardiopulmonary morbidity and mortality is one example that was presented and discussed by the presenter's meeting. The meeting participants discussed past and existing needs for air quality, air pollution exposure, and health information in the context of health effects research, surveillance activities, or tracking-oriented applications by different states. The group also identified emerging issues in acute and chronic air pollution epidemiology and surveillance, and a number of key limitations, barriers, and challenges facing many researchers dealing with access and utilization of individual-level health information.

This paper addresses these concerns by providing a synthesis of lessons learned from past air pollution health effects or public health-tracking studies on particulate matter (PM),

1. All address correspondence to: Dr. Mark J. Utell, Pulmonary Division, Department of Medicine and Environmental Medicine, University of Rochester Medical Center, Box 692, 601 Elmwood Avenue, NY 14642, USA. Tel.: +1 585 275 4861. Fax: +1 585 273 1058.
E-mail: mark_utell@umrc.rochester.edu
Received 31 October 2007; accepted 19 June 2008; published online 10 September 2008

ozone (O_3), and air toxics, and examines the particular health information needs of the health research community. Several limitations are highlighted regarding the use of routinely collected individual health data that need to be addressed to conduct high quality accountability research assessing the success of environmental control policies. In addition, this paper discusses the need for "less conventional" or emerging applications of health data sets, for the study of neurological diseases, autoimmune diseases (e.g. multiple sclerosis (MS), amyotrophic lateral sclerosis), birth outcomes, and so on. This paper concludes by identifying future opportunities for advancing the current knowledge through epidemiological accountability and tracking studies. This will ultimately lead to more informed decisions to prevent potential adverse health effects resulting from exposures to ambient air pollution.

This paper is one of three summarizing the findings and recommendations of the symposium. A summary paper (Özkaynak et al., 2008) describes the goals and issues identified by the speakers and participants, and presents the symposium recommendations. A companion paper (McKone et al., 2008) describes the usefulness and limitations of air quality data used to estimate exposure in epidemiology research and accountability studies, surveillance, and tracking activities. Taken together, the set of papers attempts to provide a background for understanding how health and air quality data have been used by public health researchers and practitioners and what improvements will be necessary to adequately accommodate new demands on these data sets for environmental health research, public health tracking, and accountability studies. For additional information on specific topics or presentations, the reader may refer to the complete copies of symposium presentations or the break-out group reports that are available from the symposium website http://www.epa.gov/nerl/symposium/.

Application of health databases to PM-associated cardiopulmonary health effects: a successful example

The example of how PM concentration in ambient air was identified as a risk factor for cardiopulmonary mortality and morbidity illustrates the historic value of administrative data sets to public health protection. Table 1 presents a brief outline of important epidemiological studies of the health effects of fine particulate and related air pollution, and the type of data used by each study.

Effects of Short-Term Exposure

Clearly, epidemiological studies of short-term effects of PM benefited greatly from the use of readily available administrative data such as mortality and hospitalization records, rather than through designed epidemiological studies. For example, short-term mortality effects ($\sim$3 to 4-fold excess) in the 1952 London smog episode were observed through the routine administrative collection of death records (UK Ministry of Health, 1953). The London smog episode data as well as the London data from later, non-episodic periods have been analyzed by various US researchers (e.g. Ware et al., 1981; Schwartz and Marcus, 1990; Bell and Davis, 2001) and provided valuable information for regulatory policy decisions. Likewise, studies using administrative data (mostly mortality records) combined with the air pollution data from major US cities such as New York from the 1970s reported associations between PM and mortality (e.g. Schimmel and Greenburg, 1972). These studies collectively laid out a framework for more modern time-series studies.

The reemergence of research interest in PM effects and the increase in the number of time-series studies in the 1990s started with a group of reports arising out of the Health Effects of Airborne Particles Center at the Harvard University Kennedy School of Government's Energy and Environmental Policy Center during the 1980s (Özkaynak et al., 1983). These reports led to the subsequent publication of some of the earliest US studies relating metrics of fine particle pollution and mortality (e.g. Kinney and Özkaynak, 1991). This activity was then followed by a series of studies published in the 1990s finding associations between daily morbidity/mortality and PM whose levels were mostly well below the prevailing ambient air quality standards (e.g. Pope, 1989; Pope et al., 1992; Schwartz and Dockery, 1992). Recently, there have been several large multicity daily time-series studies that have reduced concerns about city selection and publication bias (Samet et al., 2000a, b; Dominici et al., 2003a, b; Analitis et al., 2006).

Early studies of the effects of "summer haze" pollution on respiratory morbidity (e.g. Thurston et al., 1992) examined the relationship between air pollution and hospital admissions. There also have been several multicity studies of hospitalizations (Schwartz, 1999; Samet et al., 2000a, b; Dominici et al., 2006). In addition, a methodological alternative to the time-series analysis, the case-crossover design has been used to evaluate the effects of short-term exposure to various health end points including mortality, hospitalization, and acute cardiovascular events (Pope et al., 2006). Epidemiologic studies of effects of short-term PM exposure on mortality and hospitalizations have relied primarily on the availability of administrative data. However, as indicated in Table 1, study-specific prospective collection of data and clinically collected data have also played a role in understanding the effects of short-term exposure on lung function, respiratory symptoms, markers of inflammation, or various markers of cardiac risk.

Effects of Long-Term Exposures

Although the studies of short-term exposure to air pollution strongly suggest health effects, they provide little information about the role of pollution in inducing or accelerating the progress of chronic disease and the degree of life shortening.

Table 1. Summary of epidemiological studies of the health effects of particulate air pollution.

Health end points	Primary methods	Primary sources of health data	Citations of examples
Studies of short-term exposure			
Mortality counts	Episode analysis, daily time series, case crossover	Routinely collected vital statistics	Dominici et al. (2003a, b); Ostro et al. (2006)
Hospitalizations	Episode analysis, daily time series, case crossover	Routinely collected hospitalization records	Schwartz (1999); Samet et al. (2000a, b); Dominici et al. (2006)
Infant mortality	Daily time series	Routinely collected vital statistics	Ha et al. (2003)
Lung function/respiratory symptoms	Panel-based time series	Study-specific prospective collection	Pope et al. (1991); Hoek et al. (1998)
Indicators of pulmonary/ systemic inflammation or cardiac risk	Panel-based episode or time series	Study-specific prospective collection	Peters et al. (2001a, b)
Acute cardiovascular events	Time series or case crossover	Hospitalization records and/or CV registries	Zanobetti and Schwartz (2005); Pope et al. (2006)
Cardiac arrhythmias	Case crossover	Clinically collected ICD and other data	Dockery et al. (2005)
Studies of long-term exposure			
Mortality rates	Cross-sectional regression	Mortality data	Lave and Seskin (1970); Özkaynak and Thurston (1987)
Survival data	Survival analysis, proportional hazard regression	Prospective collection of study-specific health, survival, and air pollution data or data collected separately for other purposes.	Dockery et al. (1993); Pope et al. (1995, 2002); Krewski et al. (2000); Laden et al. (2006); Lipfert et al. (2006); Jerrett et al. (2005); Miller et al. (2007)
Infant mortality	Cross-sectional	Linked birth and mortality data	Woodruff et al. (1997, 2006)
Lung function, CIMT, and other measures of disease	Cross-sectional or long-term longitudinal	Study-specific collected data or links to other data sets	Dockery et al. (1996); Gauderman et al. (2004, 2007); Künzli et al. (2005)
Natural experiment or intervention studies			
Mortality	Descriptive comparative statistics, regression analysis	Routinely collected vital statistics	Clancey et al. (2002); Hedley et al. (2002); Pope et al. (2007)
Hospitalizations	Descriptive comparative statistics, regression analysis	Routinely collected hospitalization records	Pope (1989, 1991); Friedman et al. (2001)
Infant mortality	Econometric methods	Routinely collected vital statistics	Chay and Greenstone (2003a, b)

Abbreviations: CIMT, carotid intimal medial thickness; CV, cardiovascular; ICD, implantable cardioverter defibrillator.

There have been various research efforts to explore the effects of longer term exposures including: (1) population-based, cross-sectional studies of mortality rates; (2) cohort-based mortality studies; and (3) panel- or cohort-based studies of long-term exposure on measures of disease, lung function, subclinical atherosclerosis, and so on.

Lave and Seskin (1970) reported an early population-based, cross-sectional mortality analysis that found that mortality rates across US metropolitan areas were associated with long-term average air pollution, especially fine PM air pollution. Similar studies by other researchers have also linked long-term fine particulate air pollution exposure to increased mortality rates (Özkaynak and Thurston, 1987). Although these studies attempted to control for population average differences in demographic, socioeconomic, and various other ecologic variables, their findings stimulated interest in conducting prospective studies that could control for individual risk factors such as cigarette smoking. These earlier studies relied almost entirely on readily available administrative data, mostly mortality and census records.

Two early prospective cohort studies, the Harvard Six Cities Study (Dockery et al., 1993) and the American Cancer Society (ACS) study (Pope et al., 1995) reported increased risks of cardiopulmonary mortality associated with spatial differences in long-term average concentrations of fine particulate air pollution even after controlling for cigarette smoking and numerous other individual risk factors. The Harvard Six Cities Study involved the prospective collection of study-specific health, survival, and air pollution data. The ACS study and subsequent cohort-based mortality studies (Lipfert et al., 2006; Miller et al., 2007) generally have not involved prospective collection of study-specific air pollution and health data, but have relied on health, survival, and risk factor data collected separately for other purposes. Extensive reanalysis (Krewski et al., 2000) and extended analyses (Pope et al., 2002; Jerrett et al., 2005; Laden et al., 2006) of the Harvard Six Cities and ACS cohort studies demonstrated that the fine PM mortality risk estimates were remarkably robust to various statistical modeling specifications and to control for individual risk factors, including age, sex, race, smoking, alcohol use, marital status, education, body mass, occupational exposures, and diet or the inclusion of spatial smoothing (Pope et al., 2002) or ecological socioeconomic variables (Krewski et al., 2000; Jerrett et al., 2005). Recently, Miller et al. (2007) linked air pollution data with risk factor, health, and survival data from the Women's Health Initiative clinical trial and observational study. They reported strong associations between long-term exposure to fine particulate air pollution and cardiovascular disease and death among postmenopausal women. Although these studies obtained individual-level information about their cohort subjects, they relied upon state and national death registries for the determination of death and its cause(s). These cohort studies were also consistent with, and largely confirmed the results from the earlier cross-sectional studies discussed above, that relied solely on administrative death records and publicly available air pollution exposure data.

As the link between mortality and air pollution was established, several studies have reported that long-term exposure to air pollution is associated with health outcomes not generally measured as part of routine health surveillance activities. These outcomes include deficits in lung function (Schwartz, 1989; Chestnut et al., 1991; Raizenne et al., 1996) and increased chronic cough, bronchitis, and chest illness (Schwartz, 1993; Dockery et al., 1996; Peters et al., 1999). Most of these studies have involved collecting study-specific data sets, but Chestnut et al. (1991) and Schwartz (1989) used data from the National Health and Nutrition Examination Survey. A notable recent study of the effects of long-term air pollution exposure is the Southern California Children's Health Study that involved study-specific prospective collection of lung function and other health data on thousands of children living in 12 southern California communities with differing long-term average concentrations of air pollution. Significant deficits in 8-year lung function growth (from approximately 10–18 years of age) was observed for children living in more polluted communities (Gauderman et al., 2004) and for children living near major roadways (Gauderman et al., 2007). Künzli et al. (2005) reported another interesting study of effects of long-term air pollution exposure in the Los Angeles metropolitan area. They linked health data from two clinical trials with pollution data and found that long-term fine particle exposure was associated with carotid intima-media thickness, a measure of subclinical atherosclerosis.

This abbreviated discussion of the literature on cardiopulmonary health effects and PM is meant to be an example of the successful use of routine administrative data to identify and evaluate environmentally related chronic disease. One major advantage of analyzing administrative health data for environmental impact assessment is that, because the data are often collected in a uniform manner on a statewide basis (e.g. hospital admissions) or nationwide basis (e.g. mortality), they can be linked to the nationwide air pollution database and can be analyzed with a consistent approach across different cities. In addition, as they are routinely collected, they can be analyzed much more cost-effectively than in investigator-designed studies where both the environmental and health data must be collected by the investigators. For example, the National Mortality and Morbidity Air Pollution Studies (NMMAPS) and studies using Medicare hospital admissions data have used these data sets to conduct powerful analyses of health effects across many US cities (Samet et al., 2000a, b; Dominici et al., 2003a, b; Table 1). These analyses provided summary risk estimates with greater statistical power and precision than possible from single-city studies. The NMMAPS project also led to the Internet-Based Health and Air Pollution Surveillance System (iHAPSS), which was developed and maintained by the Johns Hopkins University and currently provides mortality (daily counts), air pollution, and meteorological data for 108 US cities, as well as statistical models (code) that the original researchers used in their analyses. Thus, the iHAPSS database allows researchers to not only reproduce the original findings but also apply their own models. The iHAPPS provided a model framework to combine administrative data and transparent research tools without risking violation of privacy (because only the aggregated data are released). Although the use of administrative data sets has been advantageous for health research, currently they are not an ideal source for timely, high-quality data needed for epidemiological studies of environmental–health interactions.

Challenges of administrative health data: standardization, privacy, linkage, and accountability

Although administrative health record data have clearly provided great opportunities for epidemiological studies of short- and long-term PM (and other pollutants') health effects, there are limitations associated with these data, some of which are inherent, and others are policy related. Health research and surveillance in the United States is presently conducted within a decentralized jurisdictional framework for data collection, compilation, and distribution. Responsibility for vital statistics records lies with the states who report these data to the National Center for Health Statistics (NCHS) at the CDS. Other records, including patient hospital or emergency room admissions, doctor's office visit records, insurance reimbursement records, and others are maintained by the organizations that develop them. Although some states do have in place mandatory reporting requirements for hospitalizations, centralized systems for collecting outpatient (including emergency room) services or pharmaceutical purchases are generally non-existent.

The lack of standardization and completeness of non-Medicare health records collected by disparate health providers has presented a significant barrier to the conduct of health tracking or epidemiology studies of air pollution using daily health data across the many cities and states of the United States. Individual states in the United States have handled data collection and release differently, with individual record details not always available, and with data often being provided to researchers in an aggregated form, such as into daily counts by county. Such data-release limitations, whereas protecting the privacy of patients, have restricted the ability of researchers to assign individual exposure estimates at smaller geographic units, or to investigate potential air pollution effect modifiers, such as race, ethnicity, health insurance coverage, and socioeconomic status.

Roberts and colleagues (2006b) documented the difficulties in assembling and standardizing rudimentary asthma-related event data among even a small number of providers. Particularly among capitated health plans, for which reimbursement is not calculated on the basis of individual units of service provided, there is little incentive for routine, timely, or complete reporting of health service events. Several commentators (Ungar et al., 2002; Lanphear and Gergen, 2003; Centers for Disease Control, 2005) have outlined the need for multiple outcome measures to become part of disease surveillance, particularly for asthma. The development of registries for relevant diseases would resolve some of the issues. The benefits of such registries for these purposes have been previously discussed in testimony before the US Senate (Thurston, 2002) relative to proposed US legislation to promote the development of nationwide health effects tracking registries, such as Bill S1442, entitled: "Coordinated Environmental Health Network Act of 2005" (US Senate, 2005).

Over the last 10 years, national statutes, including the Health Insurance Portability and Accountability Act (HIPAA) of 1996 and the Family Educational Rights and Privacy Act (FERPA) of 1974, as well as additional state laws, were enacted to protect the privacy of US citizens and to guide health providers and other organizations concerning data collection, compilation, and reporting of individual records. HIPAA was originally enacted by the US Congress in 1996, but various rules, including the Privacy Rule, were created and took effect in more recent years. Title II of HIPAA, the Administrative Simplification provisions, requires the establishment of national standards for electronic health-care transactions, but also address the security and privacy of health data. These rules also affected the policies of academic institutions, including mandatory HIPAA training courses and regulations regarding the Internet and computer usage within institutions. The Internal Review Boards (IRBs) of academic institutions/medical schools have also incorporated these HIPAA rules. All research projects need to obtain an IRB approval before they can start, even if the study involves analysis of secondary data, as in the case of mortality and morbidity analyses (considered as "exempt" from full human subject protocol reviews under National Institute of Health guidelines). Thus, these IRB-approved studies are in effect complying with the HIPAA rules, which are in fact positive changes to ensure protecting the subjects' privacy.

Although HIPAA plays a role in the accessibility of health data for research on hospitalization and outpatient events, FERPA has arisen as a parallel issue in the field of neurodevelopment and behavioral disorders. Similar to the purpose of HIPAA in the context of health-care provision, FERPA exists as a measure to safeguard records in schools, as well as to ensure access to such records by parents; unlike HIPAA, however, FERPA contains no exclusionary language to allow for public health-related activities, even when the records in question are health related in content. Most notably this has become evident as a conflict with respect to the study of autism, although the study of any disabling condition relevant to the educational needs of children could be affected. Although the precise consequences of FERPA for public health work were long in question, the Federal Department of Education declared in January 2003 that any sharing of records for this purpose ran counter to the statute, causing the disruption, and even cessation of several CDC-funded activities to assess the magnitude and nature of autism as a public health issue. Until this interpretation of the statute is renegotiated or the legislation itself is revised, FERPA is likely to represent a major barrier to progress in the understanding of the epidemiology of autism and other neurodevelopmental disorders.

Privacy concerns have led the states to impose more restrictions on access to the vital statistics they provide to the NCHS. Until 2004, epidemiologists and other researchers signed data protection agreements with the US CDS to obtain vital statistics data provided by state health departments and compiled and released by the NCHS for research purposes. In response to concerns raised by some states, NCHS issued a new data use policy that significantly restricted access to these data for research. Since that time, academic researchers have been presented with two alternatives, both financially burdensome and administratively arduous, which have resulted in the abrupt halt of studies evaluating the association of daily changes in air pollution and mortality for the United States population. Under the NCHS data policy, health investigators can conduct their analyses using a custom data set containing the necessary variables kept at the NCHS Research Data Center. The data set may not leave the data center and a researcher must travel to the Center to conduct analyses or have NCHS staff conduct the analyses and send the output electronically. Remote analyses may be possible under some circumstances. Alternatively, a researcher is faced with the task of requesting release of the data from each of the more than 50 states and other jurisdictions in the United States. Obviously this presents a financial and administrative barrier for anyone contemplating analyses in more than a few jurisdictions.

Interestingly, unlike the NCHS, the Center for Medicare and Medicaid Services (CMS), whose policy would be more directly impacted by the HIPAA rules, has not stopped releasing the hospital admissions data for academic research projects. As the CMS requires a full review of study protocols and issues a detailed data use agreement before researchers can obtain the data, there appears to be enough protection of the subjects' privacy. Given the apparent efficacy of the more "research friendly" CMS data-release policy, the NCHS should consider revising their data-release policies to be more consistent with that of the CMS.

Another limitation in present-day health databases is the lack of linkage in health records across databases, from doctor's visits, emergency department visits, hospital admissions, and death records. With this information, it is potentially possible to investigate the coherence of effects leading to mortality from air pollution. For example, Canada has national health care, and such records are available to researchers to investigate the relationship between morbidity and mortality from environmental factors. Indeed, Goldberg et al. (2001) used such Canadian health-care record data (including medication use) to identify sub-populations in Montreal, Canada, that were at especially elevated risk of death from acute exposure to ambient air pollution. Such a linking of US morbidity and mortality health records by individual, and making them available to health researchers, would lead to a greatly enhanced ability to identify especially susceptible sub-populations for the adverse health effects of air pollution.

Protecting public health from environmental risks generally requires regulatory actions and tracking their consequences to determine their effectiveness. Accountability in this symposium referred to the extent to which either a regulatory or natural intervention can be shown to have influenced public health outcomes. Research on accountability is in its early stages, but even now significant challenges in assessing the health impacts can be expected. A Health Effects Institute Report (2003) laid out a chain of accountability, which characterizes links between regulatory action and health effects. Following a regulatory intervention, the chain includes a reduction in emissions and improved ambient air quality that is then translated into a reduced exposure or dose and the reduction of an adverse health effect.

Several well-defined natural experiments or interventions that evaluated abrupt changes in pollution exposure (over a period of several months to a year or so) provide the best available evidence to date for documentation of the health benefits following reductions in air pollution levels (Table 2). Although these studies cannot evaluate changes in exposure over years or decades, they explore effects of much longer exposure than the short episode, daily times series and related studies. For example, the abrupt, 13-month reduction in air pollution that accompanied the steel mill closure in Utah Valley was associated with reduced respiratory hospitalizations and cardiopulmonary mortality (Pope, 1996). Substantial long-term reductions in air pollution resulting from the ban on the use of coal burning in Dublin, Ireland, were associated with reductions in respiratory and cardiovascular deaths (Clancy et al., 2002). Reductions in seasonal mortality were observed in the first 12 months after restrictions on sulfur content of fuel in Hong Kong (Hedley et al., 2002). Chay and Greenstone (2003a, b) utilized recent econometric tools (including regressions with instrumental variables, and regression discontinuity) to explore the quasi-natural experiments associated with new air pollution standards and an economic recession. Finally, a recent study retrospectively analyzed historic records related to an $8\frac{1}{2}$-month copper smelter strike in the United States in the late 1960s (Pope et al., 2007). Strike-related sharp reductions in sulfate PM and related air pollutants across four southwestern states (New Mexico, Arizona, Utah and Nevada) were associated with small but significant reductions in mortality even after controlling for time trends, mortality in bordering states, and nationwide mortality counts for influenza/pneumonia, cardiovascular, and other respiratory deaths. All of these natural experiments or intervention studies relied primarily on available administrative data, again demonstrating that, to evaluate the effectiveness and merits of environmental control measures, careful documentation of health, such as through the availability of administrative and health registry data, will be required for optimum health policy implementation.

The accountability studies discussed above were conducted with data from different countries and innovatively examined the health benefits associated with sharp increases in local air pollution levels as a result of either planned regulatory actions or unexpected disruptions in industrial operations. Unfortunately, studying the impacts of more subtle changes of pollution, in particular complex mixtures of gases and particles with varying toxicities pose considerable challenges in detecting health risks attributable to specific pollutants or regulatory actions. Recently, epidemiologists (Burnett et al., 2005; Laden et al., 2006; Dominici et al., 2007; Jerrett et al., 2007) have begun analyzing changes in the particulate mortality risk coefficients or effects in United States and Canada as a result of changing ambient pollution levels over the past decade or more. Typically, however, significant reductions in pollution levels or changes in composition of pollutants of concern take place over a long period of time further complicating the collection and analysis of air quality and health data for the purposes of accountability assessments (Tolbert, 2007). Clearly, this is an area where further research and data are definitely needed. In particular, consistent and continuous collection of health and air quality

Table 2. Summary of accountability studies of the health effects of particulate air pollution.

Study	Measure or event resulting in improved air quality	Health end points	Primary method	Primary source of health data
Clancy et al. (2002)	Coal ban in Dublin	Age-standardized total and cause-specific death rates	Interrupted time-series analysis, adjusting for weather, respiratory epidemics, and death rates in the rest of Ireland, comparing death rates 72 months before and after the ban on coal sales in Dublin	Central Statistics Office
Friedman et al. (2001)	Reduced traffic during Atlanta Olympics	Citywide acute care visits and hospitalizations for asthma (asthma events) and non-asthma events	Ecological study comparing the 17 days of the Olympic Games (19 July–4 August 1996) to a baseline period consisting of the 4 weeks before and 4 weeks after the Olympic Games	Georgia's Medicaid claims file, the patient database of a HMO, computerized emergency department records from two of the three pediatric hospitals in Atlanta, and the Georgia Hospital Discharge Database
Hedley et al. (2002)	Low sulfur fuel in Hong Kong	Age-stratified monthly death rates (total and cause specific)	Poisson regression comparing 5-year periods before and after fuel change, adjusting for time trend, seasonality (sine and cosine terms), temperature, and humidity	Census and Statistics Department, Hong Kong Special Administrative Region. 1985–1995 known death micro-data sets (accessed 5 June 2002)
Heinrich et al. (1999, 2000, 2002)	Improved air quality in association with German reunification	Community-specific adjusted prevalence of non-allergic respiratory symptoms	Three cross-sectional surveys of children, age 5–14 years, in three former East German metropolitan areas during the decade following German reunification. These surveys were conducted in 1992–1993, 1995–1996, and 1998–1999	Study-specific questionnaires were completed by parents
Pope (1989)	Closure of Utah Valley steel mill	In-patient hospital admissions data for respiratory-related illnesses between April 1985 and February 1988 (compiled monthly)	Regression models assessing the association between hospital admissions and particulate pollution (PM10) in Utah Valley during the period April 1985–February 1988. This time period included the closure and reopening of the local steel mill, the primary source of PM10	Inpatient admissions data were collected from three hospitals
Tonne et al. (2008)	CCS introduced in February, 2003, to alleviate traffic congestion in Central London	All-cause mortality	Concentrations of NO_2 and PM_{10} were estimated using a combined emission dispersion and regression modeling approach for areas within and outside the charging zone. Years of life gained among residents in the two areas, stratified by SES, were compared in a quantitative risk assessment using mortality–pollution relationships derived from published epidemiology studies	Office for National Statistics. Deaths—England and Wales

Abbreviations: CCS, Congestion Charging Scheme; HMO, health maintenance organization; SES, socioeconomic status.

data for PM and co-pollutants over multiple sites in many urban locations will be necessary for being able to determine the health benefits associated with different air quality management programs or regulations.

Emerging applications of health data sets

Although the challenges of access to information in administrative data sets are many, the expansion of their utility begins to emerge. Public health practitioners and regulators continue to require access to vital statistics and health-care records to track trends in health status in relation to identified air quality risks, identify new environmental diseases, and to determine the success of risk reduction measures. Although the utility of these data sets has been clearly demonstrated for linking cardiopulmonary diseases to specific air pollutants, their expanded utility for less traditional applications warrants further discussion. We address below a number of these emerging areas of concern for use of health information in future air pollution health investigations.

Applying Health Databases beyond Cardiopulmonary Epidemiology

What if a valid link between an environmental exposure and a specific disease exists, but has not been established? Administrative and health registry data sets might be used to track diseases suspected to have an environmental association; however, this approach presents several difficulties at this point in time. Limiting factors include: difficulties identifying all relevant cases of the health effect, limitations in estimating an individual's exposure to an environmental hazard, and delineating a biologically plausible environmental–health association. As each component of the analysis process has its own issues, the challenge is further compounded by the limitations inherent in merging health and environmental information. Although these complexities are tolerable for a number of environmental health issues, they are less manageable for many of the "emerging diseases." To better demonstrate these issues, a direct comparison of three health outcomes was completed (Table 3). Preterm birth, asthma, and MS are diseases that cover a spectrum of uncertainty concerning the reliability of case ascertainment, the degree of temporal and spatial resolution in the data to allow a reasonable estimate of exposure, and the scientific basis for linking a specific contaminant with the disease. Preterm birth represents what is currently a best case scenario, because it meets, or is close to meeting many of the requisite conditions, although it is not ideal. Asthma rates well for some of the requisite conditions, particularly, because it has a strong scientific association with ambient air contaminants. Of all these health effects, the most challenging can be exemplified by the neurological disease, MS, and the major obstacles are defined here.

The prevalence of MS in the United States is estimated to be 250,000–350,000 cases (Anderson et al., 1992; McDonald et al., 2001). The issue of case ascertainment is significant for chronic diseases such as MS where cases cannot be readily identified through traditional administrative data sets, and standardized national or international registries, such as those for cancer, do not exist. This situation is likely to improve over time, as there are strong efforts to develop registries, such as the North American Committee on Multiple Sclerosis Patient Registry (NARCOMS).

The development of reliable registries is challenged by the issue that standardized diagnosis has been historically problematic and often inconsistent over time or from one location to another (Uitdehaag et al., 2005). Science and experience are improving our ability to more definitively identify MS cases (Poser et al., 1983; McDonald et al., 2001), but much remains to be done to achieve widespread,

Table 3. Characteristics of data required for analyzing the linkage between environmental exposures and specific health outcomes.

Requirement	Examples of health outcomes[a]		
	Preterm birth	Asthma	Multiple sclerosis
Requirements related to health tracking			
Data that are collected for the purpose of linkages needed for environmental health tracking, accountability, and research. Data are collected using a standardized protocol and compiled in a centralized administrative system with linkage capability	+ + + +	+ +	+
Case definitions that are consistent cross-sectionally and longitudinally	+ + + + +	+	+ + + (+ +)
Database/registries capture all cases in a timely manner. All cases within geographic boundaries are ascertained completely and in a timely manner	+ + + + +	+ +	+
High spatial resolution for locations of cases with residential address (preferably locations where potentially relevant exposures occur)	+ + + + +	+ +	+
Availability of covariate data collected by the same administrative system or through linkage by identification number to account for other factors such as diet, genetics, etc.	+ + + +	+ +	+
Requirements related to linkage with potentially relevant hazard data			
Databases that provide information regarding the location and time of exposures potentially relevant to the outcome	+ + +	+ + +	+
Consistent and matching temporal and geographical resolutions among potentially relevant environmental data sets	+ + +	+ + +	+
Potential for accurate measurement of individuals' exposures to potentially relevant contaminants	+ +	+ +	+
High spatial resolution for locations of potentially relevant hazards	+ +	+ +	+
Hypothesis generation and validity issues			
Pollutants of concern have been identified	+	+ + + + +	+
Valid animal models exist for the health effect in question	+ + +	+ + +	+ +
Longitudinal studies have demonstrated the cumulative consequences of exposure over time	+ + +	+ + +	+
Latency between putative exposures and health outcomes is relatively short or methods exist to account for long latencies	+ + + +	+ + +	+

[a]Note: the classification range (+) to (+ + + + +) represents the degree of availability, quality, and scientific utility of data in terms of poor (+), fair (+ +), modest (+ + +), good (+ + + +), excellent (+ + + + +) classification. For instance, (+) indicates data generally not available and/or requisite conditions not met and (+ + + + +) indicates generally available data and conditions favorable for scientific advancement.

long-term consistency that equals that of diseases such as cancer. This state of flux is yet another obstacle to quality case ascertainment as it is typically invalid to combine cases from different periods of time when the case definition endures significant changes, or from one place to another where diagnosis habits and biases are substantially different. Most health issues have emerged from these limitations over time, and it is likely neurological diseases are on the threshold of doing the same (Polman et al., 2005). However, the current status of case ascertainment would be less than ideal for effectively and routinely linking them with environmental hazards and exposures for surveillance purposes. Additionally, without being able to do such linkages on a large scale it would be difficult to use these data either for accountability initiatives or in epidemiological research.

As efforts move toward improved case ascertainment, another challenge will be in identifying likely environmental hazards to link to the disease. MS is characterized by sensory and motor deficits associated with central nervous system demyelination. Although the neuronal changes associated with the disease have been described, the pathophysiology remains unknown. A variety of hypotheses exist about the cause of MS; however, there is general agreement that it is the product of complex interactions between environmental exposures and genetic susceptibility either to autoimmune reactions or to the negative impact of the specific exposure (Vella, 1984; Hogancamp et al., 1997; Kahana, 2000; Willer and Ebers, 2000). The most widely accepted view is that the demyelination results from immune system activities and inflammatory processes that result in an autoimmune response.

These immune system reactions are precursors to the initiation of progressive demyelination, and are the focus of hypotheses regarding environmental exposures. A popular hypothesis is that certain viral infections may initiate the cascade of detrimental responses in the central nervous system. For example, there is some evidence of an association between MS incidence and increased incidence of influenza (Kazmierski et al., 2004). Additionally, there is evidence to suggest that changes in air quality modulate individual susceptibility to such viruses in an indirect manner such that increased levels of PM exacerbate systemic immune and inflammatory responses, consequently increasing susceptibility to infectious viruses (Thomas and Zelikoff, 1999; Zelikoff et al., 2003). In addition, MS relapses have been associated with increased levels of ambient PM (Oikonen et al., 2003). These early hypotheses suggest future analytical linkages between ambient air quality and MS are warranted and necessary to better understand the impact of air pollution on many facets of human health.

Epidemiological research to address environmental–health linkages for air quality and MS is difficult using traditional epidemiological strategies primarily because data to help narrow the focus to a few, biologically plausible candidate environmental hazards are insufficient, and the latency between exposure and diagnosis is largely unknown thus complicating efforts to characterize the related exposure. Different methods can be used to match an individual to an environmental hazard both in space and time. If the health effect were relatively acute, such as a matter of days as is the case for myocardial infarctions and fine PM (Peters et al., 2001a, b; Pope et al., 2006), it would be easier to identify the location and time when the exposure occurred and calculate an estimate. In contrast, it is much more difficult to characterize exposures in persons with chronic issues, such as these emerging neurological diseases. The latency between exposure and diagnosis of these diseases, whereas largely unknown, is likely of the order of months and years as compared to days. Indeed, it is likely the exposures occur over a long period of time, and methods for estimating cumulative exposure are therefore required.

The uncertain latency between exposure and diagnosis also presents a significant obstacle to using diseases such as MS to evaluate the impact of environmental regulatory actions. Even if a regulatory change significantly reduces the incidence of the disease, this effect may not be apparent for a number of years after the environmental hazard and human exposure is reduced. A worst case scenario would be for diseases typically diagnosed in late adulthood, where there is the potential for latency on the order of decades. This argues the need for long-term surveillance using indicators of linked environmental and health data to detect changes as they occur over time.

Clearly, emerging neurological diseases such as MS are extremely important environmental health topics. However, progress needs to be made to consistently identify cases and reasonably estimate exposure to environmental hazards before hypotheses about disease causation can be explored. Programs such as the CDC Environmental Public Health Tracking effort may facilitate epidemiological research in the future by providing a mechanism for linking health and the environment over time. In the meantime, focusing on promising research on genomics and developing better information on biomarkers of exposure and effects would also improve the scientific basis of future environmental regulations.

Environmental Public Health Tracking

Environmental public health tracking represents a systematic effort by CDC to enable disparate public agencies to assemble standardized data sets that together can be used for both surveillance and environmental epidemiology. EPHT focuses on both health and environmental data sources; for the present discussion, we will highlight activities related to the enhancement of health data utility, including the use of such data for environmental epidemiology research. The obstacles identified above—including the cross-cutting needs for data standardization, confidentiality, and linkability to pollution data—are addressed through the development of data enhancement services and the stimulus of planning a nationwide EPHT network to which states can contribute.

Environmental public health tracking springs simultaneously from the crisis of the "environmental health information gap" and the possibilities of the information age (McGeehin et al., 2004). By focusing on the development of health data itself as a common good, the aims of EPHT are to simultaneously maximize its utility for multiple audiences and facilitate information dissemination (California Policy Research Center, 2004; Ritz et al., 2005). As the emergence of the concept as a focus within government and academia, tracking has served as a forum for addressing issues of data sharing and utility (Environmental Health Tracking Project Team, 2000).

As discussed previously, confidentiality rules and protocols are inconsistently applied across data systems, even when the primary uses of data systems have been identified as including research. One strategy implicit in the development of an EPHT network is that data flow can become more transparent, so that the availability of specific fields, levels of aggregation, and security protocols can be agreed upon up front. Although not a panacea for inconsistencies of confidentiality practices between agencies (both public and private), the provision of a forum in which public health information needs can receive primary attention can foster greater collaboration in this arena.

Some of the most concrete benefits of the EPHT enterprise promise to be those pertaining to issues of linkability, however. For example, a primary requirement for linkability is data standardization, so EPHT resources have focused on both the elaboration of standards and on enabling data systems to adopt these standards at little or no cost to them. Standards include architecture requirements described by the CDC's Public Health Information Network and the US Environmental Protection Agency's National Environmental Information Exchange Network (NEIEN) initiatives. Less well known are standards that enable the flow of geographic information and allow for transformation between coordinate systems and spatial intersection, which have been developed by the Open Geospatial Consortium, enabling a maximum degree of vendor neutrality and system interoperability (Percivall, 2003). Grantees of the CDC's EPHT Program have similarly developed generic specifications for calculating metrics that estimate the spatial and temporal relationship between health and hazard events.

Of course, standardization applies not only to data formatting but also to case ascertainment and definition. Among the main objectives of any health surveillance program are the comparison across municipalities, states, and time periods for the understanding of trends (Teutsch, 2000). Common definitions and methods of collection must be established for this to take place; although this may be taken as basic practice in epidemiology, there is no centralized decision-making authority for data collection or standardization that operates across the jurisdictions of governmental agencies or between states. The formulation of approaches to governance and priority setting among diverse data-generating agencies, both public and private, has been raised as a primary function for EPHT (Kyle et al., 2006).

Another component that augments data linkability is the maximization of spatial and temporal resolution for both health and hazard data (Centers for Disease Control, 2004). Highly aggregated data (for example, state- or province-wide counts) are rarely useful in environmental epidemiology; in most of the work cited previously, the data were resolved at least to the level of cities, if not further. Hazard and exposure data are rarely resolved to the level of the individual, however, because of limited resources for measurement and population mobility. The key is that the resolution of health data should at least match that of hazard data. As the science of exposure estimation progresses, the scalability of health data to a variety of levels will become more crucial.

Once health data can be expressed as points in space and time, the potential for linkage expands considerably. For example, with a geocoded data set describing hospitalizations due to asthma, records could be assigned values representing proximity of patient residence to roadways or point sources of pollution, or measured ambient concentrations of contaminants. The integration of health and hazard data in this way requires the consideration of a host of factors that may threaten scientific validity or complicate interpretation (Mather et al., 2004). These include (to name a few), the postulation of lag periods between exposure and outcome, the availability of covariates describing residential history or daily transportation patterns, and the implications of utilizing area-wide averages or spatial unit centroids when precise exposure data are not available. In this way, the question of interoperability is equally an issue of study validity as one of data capacity.

One function gaining currency in EPHT is the establishment of enterprise (centralized) address standardization, verification, and geocoding incorporating multiple reference data sets in a hierarchical manner. Geocoding has the effect of preserving the highest possible spatial resolution in a given data set. From there, aggregation to census tracts, postal codes, or other geographic units becomes straightforward using standard geographic information service tools, with the only consideration being the maintenance of large enough geographic units that the confidentiality of individual records is protected. Analogous to the spatial case, high temporal resolution of data also contributes to broad applicability, with similar restrictions on the limits of disaggregation in the interests of patient confidentiality.

Symposium findings

Symposium participants identified several lessons, needs, and future opportunities that can be derived from past experience using routinely collected health data. Routinely collected health data have proved to be essential for the identification and evaluation of health hazards in air pollution research. Epidemiology studies using death records from multiple cities on a regional and national scale have been instrumental in setting national air quality standards for PM. Unfortunately, access to state-controlled vital statistics records that have informed such pivotal research has been curtailed sharply during the last decade. This reduced access has threatened continuation of the type of research necessary to support future standard setting and the ability to evaluate regulatory and other prevention activities. Moreover, the importance of surveillance ("tracking") databases for the identification of additional health outcomes associated with air pollution exposures has been increasingly recognized as results emerge from the CDC Environmental Public Health Tracking program.

Several challenges were raised by symposium participants concerning the use of routinely collected health outcome data by epidemiologic studies and tracking programs. Such challenges threaten the validity and reliability of studies of air pollutant exposure for the US population and trends over time. Information on key variables, residence (e.g. demography and socioeconomic status, residence, diagnosis dates), in individual records, registries for additional exposure-related health outcomes (e.g. asthma, MS, diabetes, and others), and standardization of these data across jurisdictions would improve the evaluation of hypotheses linking exposure to disease, allow the tracking of trends in the nation's health status, and assist the identification of emerging health conditions related to environmental exposures (Table 3). A large federal, state, and local effort will be required to develop disease-based registries and to coordinate collection and standardization of administrative health data. New federal legislation may be required to facilitate and support health and exposure data collection, standardization, and sharing between jurisdictions.

Effort is needed by state legislatures to balance privacy concerns with the need to allow access to health data for research. Recognizing the value of the research to its environmental health protection mission, the US EPA Office of Research and Development (ORD) has attempted to assist air pollution researchers who conduct research supported by the Agency to maintain their access to vital statistics data. In 2006, ORD initiated a coordinated request on behalf of its grantees to the 50 states through the NCHS and secured permission to obtain mortality data from 48 states for 2000–2005. This effort likely will not constitute a long-term solution for the overall research community because it has been limited to only research funded by the EPA. Moreover, in spite of ORD's attempt to submit a consolidated request, several large states required individual applications from each research group, keeping the overall administrative burden high.

The importance of administrative and tracking data for accountability studies was also a focus of discussion at the symposium. To determine whether a population has benefited from a particular risk reduction action, such as regulating the composition of gasoline for automobiles, baseline health status must be established. Moreover, the data need to be collected in a standardized format over time and across jurisdictions, and contain appropriate spatial and temporal resolution. Future accountability studies will benefit from the progress made by the CDC EPHT program to establish long-term, standardized health data collection and reporting that is accessible to researchers across jurisdictions yet protects the privacy of citizens. Analogous to the need in EPHT to follow the incidence of sentinel health conditions that may be susceptible to environmental exposure to pollutants, accountability research requires the ability to monitor health responses to environmental changes that may occur relatively rapidly. To date, researchers have made use of opportunistic studies to document local changes in health status that occurred when economic decisions (e.g. an industry changes) or policies (e.g. changes in traffic patterns during the summer Olympics in a major city) were implemented that caused pollutant emissions to be reduced for a discrete time period. Registries or linked health records will allow accountability studies on a broader scale. Participants agreed that access to health records also is a current barrier for accountability studies. Confidentiality restrictions imposed by HIPAA and FERPA govern how institutions such as school systems, emergency clinics, and hospitals control the release of data for research.

In conclusion, symposium participants were in agreement that epidemiology studies of air pollution exposure in relation to health status in US cities and regions using vital statistics and other administratively collected health data have provided valuable information for state and federal public health protection initiatives, including setting standards for criteria air pollutants. Participants also agreed that current systems for the collection, compilation, linkage, and dissemination of health data across jurisdictions will not meet future needs of EPHT programs, or allow the identification of emerging environmentally related disease. A strenuous effort, perhaps including legislation such as has been previously proposed in the US Senate, will be necessary to ensure the standardization, centralization, and sharing of data sets, as well as to harmonize the interpretation of confidentiality and privacy protections across jurisdictions. These actions, combined with assuring researchers and public health practitioners appropriate access to data for evaluation of environmental risks, will result in the achievement of our environmental health protection goals.

Symposium recommendations

The following recommendations were developed from the conclusions reported by the symposium break-out group on health data needs and the discussions of the whole.

- There is a need for access to vital statistics and other health-related data at appropriate spatial and temporal resolution from all states in the US, as well as in other nations, for health research. Steps needed in the US include:
  - The federal government and states should collaborate to develop common confidentiality protections and provide for use of the data by researchers in academic institutions; interpretation of HIPAA regulations should be clarified and consistently applied across data systems.
  - Data from local health organizations should be aggregated into a centralized state system.
  - Data for each state should be compiled by the NCHS and qualified researchers should have full access to the compiled data set at either the aggregate or individual level, as necessary for the analysis, but with appropriate privacy protections.
  - FERPA legislation should be revised, or its interpretation renegotiated, to allow for the use of education records for public health research purposes as long as subject confidentiality safeguards are in place.
- The federal government and the states should collaborate to adopt a set of standardized variables to be collected for all individual health records by hospitals, health clinics, emergency rooms, and doctors' offices such as key demographic and socioeconomic variables, residence address, and dates of a health event or diagnosis in the individual record.
- Data should be available to researchers at appropriate spatial, temporal, and individual resolution, as required.
- The federal government and the states should collaborate to identify important health conditions or diseases associated with environmental exposures (e.g. asthma, autism, MS, diabetes, and others), to:
  - Develop a common definition for the condition or disease.
  - Delineate the latency between exposure and diagnosis.
  - Fund and develop state and national registries for these conditions or diseases.

Acknowledgements

We are grateful to the excellent contributions made by the symposium attendees who participated in the health breakout group discussions on the second day of the meeting.

Disclaimer

The United States Environmental Protection Agency through its Office of Research and Development partially funded and collaborated in the research described here under contract numbers: EP06D000757 to Dr. George Thurston; EP06D000748 to Dr. Arden Pope, and; EP06D000741 to Dr. Mark Utell. This paper has been subjected to Agency Review and approved for publication.

References

Analitis A., Katsouyanni K., Dimakopoulou K., Samoli E., Nikoloulopoulos A.K., Petasakis Y., Touloumi G., Schwartz J., Anderson H.R., Cambra K., Forastiere F., Zmirou D., Vonk J.M., Clancy L., Kriz B., Bobvos J., and Pekkanen J. Short-term effects of ambient particles on cardiovascular and respiratory mortality. *Epidemiology* 2006: 17: 230–233.

Anderson D.W., Ellenberg J.H., Leventhal C.M., et al. Revised estimate of the prevalence of multiple sclerosis in the United States. *Ann Neurol* 1992: 31(3): 333–336.

Bell M.L., and Davis D.L. Reassessment of the lethal London fog of 1952: novel indicators of acute and chronic consequences of acute exposure to air pollution. *Environ Health Perspect* 2001: 19: 389–394.

Burnett R.T., Bartlett S., Jessiman B., Blagden P., Samson P.R., Akmak S., Stieb D., Raizenne M., Brook J.R., and Dann T. Measuring progress in the management of ambient air quality: the case for population health. *J Toxicol Environ Health A* 2005: 68: 1289–1300.

California Policy Research Center. *Strategies for Establishing an Environmental Health Surveillance System in California*. California Policy Research Center, Office of the President, University of California, Oakland, 2004.

Centers for Disease Control. *Environmental Public Health Tracking Data Linkages Workgroup Final Working Report*. Centers for Disease Control and Prevention, US Public Health Service, Atlanta, 2004.

Centers for Disease Control. *Environmental Public Health Indicators Project*. Atlanta: Centers for Disease Control and Prevention 2005. Available: http://www.cdc.gov/nceh/indicators/[accessed December 1 2005].

Chay K.Y., and Greenstone M. *Air Quality, Infant Mortality, and the Clean Air Act of 1970*. 2003a, NBER Working Paper Series, No. 10053, National Bureau of Economic Research, Inc., USA.

Chay K.Y., and Greenstone M. The impact of air pollution on infant mortality: evidence from geographic variation in pollution shocks induced by a recession. *Quart J Econ* 2003b: 118(3): 1121–1167.

Chestnut L.G., Schwartz J., Savitz D.A., and Burchfiel C.M. Pulmonary function and ambient particulate matter: epidemiological evidence from NHANES I. *Arch Environ Health* 1991: 46(3): 135–144.

Clancy L., Goodman P., Sinclair H., and Dockery D.W. Effect of air-pollution control on death rates in Dublin, Ireland: an intervention study. *Lancet* 2002: 360: 210–1214.

Dockery D.W., Cunningham J., Damokosh A.I., Neas L.M., Spengler J.D., Koutrakis P., Ware J.H., Raizenne M., and Speizer F.E. Health effects of acid aerosols on North American children: respiratory symptoms. *Environ Health Perspect* 1996: 104: 500–505.

Dockery D.W., Luttmann-Gibson H., Rich D.Q., Link M.S., Mittleman M.A., Gold D.R., Koutrakis P., Schwartz J.D., and Verrier R.L. Association of air pollution with increased incidence of ventricular tachyarrhythmias recorded by implanted cardioverter defibrillators. *Environ Health Perspect* 2005: 113(6): 670–674.

Dockery D.W., Pope III C.A., Xu X., Spengler J.D., Ware J.H., Fay M.E., Ferris B.G., and Speizer F.A. An association between air pollution and mortality in six US cities. *N Engl J Med* 1993: 329: 1753–1759.

Dominici F., Daniels M., McDermott A., Zeger S.L., and Samet J. Shape of the exposure–response relation and mortality displacement in the NMMAPS database. In: *Revised Analyses of Time-Series of Air Pollution and Health. Special Report*. Health Effects Institute, Boston, MA, 2003a, pp. 91–96.

Dominici F., McDermott A., Daniels M., Zeger S.L., and Samet J.M. Mortality among residents of 90 cities. In: *Revised Analyses of Time-Series Studies of Air Pollution and Health. Special Report*. Health Effects Institute, Boston, MA, 2003b, pp. 9–24.

Dominici F., Peng R.D., Bell M.L., Pham L., McDermott A., Zeger S.L., and Samet J.M. Fine particulate air pollution and hospital admission for cardiovascular and respiratory disease. *J Am Med Assoc* 2006: 295: 1127–1134.

Dominici F., Peng R.E., Zeger S.L., White R.W., and Samet J.M. Particulate air pollution and mortality in the United States: did the risks change from 1987 to 2000? *Am J Epidemiol* 2007: 166: 880–888.

Environmental Health Tracking Project Team, Johns Hopkins School of Hygiene and Public Health. *"America's Environmental Health Gap: Why the Country Needs a Nationwide Health Tracking Network"*. The Pew Environmental Health Commission, Baltimore, 2000. Available at http://healthyamericans.org/reports/files/healthgap.pdf (accessed April 2007).

Friedman M.S., Powell K.E., Hutwagner L., Graham L.M., and Teague W.G. Impact of changes in transportation and commuting behaviors during the 1996 summer Olympic Games in Atlanta on air quality and childhood asthma. *JAMA* 2001: 285: 897–905.

Gauderman W.J., Avol E., Gilliland F., Vora H., Thomas D., Berhane K., McConnell R., Kuenzli N., Lurmann F., Rappaport E., Margolis H., Bates D., and Peters J. The effect of air pollution on lung development from 10 to 18 years of age. *N Engl J Med* 2004: 351: 1057–1067.

Gauderman W.J., Vora H., McConnell R., Berhane K., Gilliland F., Thomas D., Lurmann F., Avol E., Kunzli N., Jerrett M., and Peters J. Effect of exposure to traffic on lung development from 10 to 18 years of age: a cohort study. *Lancet* 2007: 369(9561): 571–577.

Goldberg M.S., Burnett R., Bailar III J.C., Tamblyn R., Ernst P., Flegel K., Brook J., Bonvalot Y., Singh R., Valois M.-F., and Vincent R. The identification of persons with cardio-respiratory conditions who are at risk of dying from the acute effects of ambient air particles. *Environ Health Perspect* 2001: 109(Suppl. 4): 487–494.

Ha E.-H., Lee J.-T., Kim H., Hong Y.-C., Lee B.-E., Park H.-S., and Christiani D.C. Infant susceptibility of mortality to air pollution in Seoul, South Korea. *Pediatrics* 2003: 111(2): 284–290.

Hedley A.J., Wong C.M., Thach T.Q., Ma S., Lam T.H., and Anderson H.R. Cardiorespiratory and all-cause mortality after restrictions on sulphur content of fuel in Hong Kong: an intervention study. *Lancet* 2002: 360: 1646–1652.

Heinrich J., Hoelscher B., Frye C., Meyer I., Pitz M., Cyrys J., Wjst M., Neas L., and Wichmann H.E. Improved air quality in reunified Germany and decreases in respiratory symptoms. *Epidemiology* 2002: 13(4): 394–401.

Heinrich J., Hoelscher B., Jacob B., Wjst M., and Wichmann H.E. Trends in allergies among children in a region of former East Germany between 1992–1993 and 1995–1996. *Eur J Med Res* 1999: 4(3): 107–113.

Heinrich J., Hoelscher B., and Wichmann H.E. Decline of ambient air pollution and respiratory symptoms in children. *Am J Respir Crit Care Med* 2000: 161(6): 1930–1936.

Hoek G., Dockery D.W., Pope III C.A., Neas L., Roemer W., and Brunekreef B. Association between PM_{10} and decrements in peak expiratory flow rates in children: reanalysis of data from five panel studies. *Eur Respir J* 1998: 11: 1307–1311.

Hogancamp W.E., Rodriguez M., and Weinshenker B.G. The epidemiology of multiple sclerosis. *Mayo Clin Proc* 1997: 72: 871–878.

Jerrett M., Burnett R.T., Ma R., Pope III C.A., Krewski D., Newbold K.B., Thurston G., Shi Y., Finkelstein N., Calle E.E., and Thun M.J. Spatial analysis of air pollution and mortality in Los Angeles. *Epidemiology* 2005: 16(6): 727–736.

Jerrett M., Newbold K., Burnett R., Thurston G., Lall R., Pope C., Ma R., De Luca P., Thun M., Calle J., and Krewski D. Geographies of uncertainty in the health benefits of air quality improvements. *Stoch Environ Res Risk Assess* 2007: 21: 511–522.

Kahana E. Epidemiologic studies of multiple sclerosis: a review. *Biomed Pharmacother* 2000: 54: 100–102.

Kazmierski R., Wender M., Guzik P., and Zielonka D. Association of influenza incidence with multiple sclerosis onset. *Folia Neuropathologica* 2004: 42: 19–23.

Kinney P.L., and Özkaynak H. Associations of daily mortality and air pollution in Los Angeles County. *Environ Res* 1991: 54: 99–120.

Krewski D., Burnett R.T., Goldberg M.S., Hoover K., Siemiatycki J., Jerrett M., Abrahamowicz M., and White W.H. *Reanalysis of the Harvard Six Cities Study and the American Cancer Society Study of Particulate Air Pollution and Mortality. Special Report*. Health Effects Institute, Cambridge, MA, 2000.

Künzli N., Jerrett M., Mack W.J., Beckerman B., LaBree L., Gilliland F., Thomas D., Peters J., and Hodis H.N. Ambient air pollution and atherosclerosis in Los Angeles. *Environ Health Perspect* 2005: 113(2): 201–206.

Kyle A., Balmes J., Buffler P., and Lee P. Integrating research, surveillance, and practice in environmental public health tracking. *Environ Health Perspect* 2006: 14(7): 980–984.

Laden F., Schwartz J., Speizer F.E., and Dockery D.W. Reduction in fine particulate air pollution and mortality: extended follow-up of the Harvard Six Cities Study. *Am J Respir Crit Care Med* 2006: 173(6): 667–672.

Lanphear B., and Gergen P. Invited Commentary: Asthma Surveillance in US Children. *Am J Epidemiol* 2003: 158(2): 105–107.

Lave L.B., and Seskin E.P. Air pollution and human health. *Science* 1970: 169: 723–733.

Lipfert F.W., Wyzga R.E., Baty J.D., and Miller J.P. Traffic density as a surrogate measure of exposures in studies of air pollution health effects: long-term mortality in a cohort of US Veterans. *Atmos Environ* 2006: 40: 154–169.

Mather F., White L., Langlois E., Shorter C., Swalm C., Shaffer J., et al. Statistical methods for linking health, exposures, and hazards. *Environ Health Perspect* 2004: 112(14): 1440–1445.

McDonald W.I., Compston A., and Edan G. Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the Diagnosis of Multiple Sclerosis. *Ann Neurol* 2001: 50: 121–127.

McGeehin M., Qualters J., and Niskar A. National environmental public health tracking program: bridging the information gap. *Environ Health Perspect* 2004: 112(14): 1409–1413.

McKone T.E., Ryan P.B., and Özkaynak H. Exposure information in environmental health research: current opportunities and future directions for particulate matter, ozone, and toxic air pollutants. *J Expos Sci Environ Epidemiol* 2008; doi:10.1038/jes.2008.3.

Miller K.A., Siscovick D.S., Sheppard L., Shepherd K., Sullivan J.H., Anderson G.L., and Kaufman J.D. Long-term exposure to air pollution and incidence of cardiovascular events in women. *N Engl J Med* 2007: 356: 447–458.

Ministry of Health (UK). *The Report of the Chief Medical Officer on the State of the Public Health*. Her Majesty's Stationery Office, London, 1953.

Oikonen M., Laaksonsn M., Laippala P., Oksaranta O., Lilius E.M., Lindgren S., et al. Ambient air quality and occurrence of multiple sclerosis relapse. *Neuroepidemiology* 2003: 22: 95–99.

Ostro B., Broadwin R., Green S., Feng W.-Y., and Lipsett M. Fine particulate air pollution and mortality in nine California counties: results from CALFINE. *Environ Health Perspect* 2006: 114(1): 29–33.

Özkaynak H., and Thurston G.D. Associations between 1980 US mortality rates and alternative measures of airborne particle concentration. *Risk Anal* 1987: 7(4): 449–461.

Özkaynak H., Glenn B., Qualters J.R., Strosnider H., McGeehin M.A., and Zenick H. Summary and findings of the EPA and CDC symposium on air pollution exposure and health. *J Exp Sci Environ Epidemiol* 2008; doi:10.1038/jes.2008.34.

Özkaynak H., Thurston G.D., Tosteson T.D., Smith C.M., Kinney P.L., Beck B., Scornik W., Colome S.D., Chow J., Schatz A., and Matthews P. Analysis of health effects resulting from population exposures to ambient particulate matter. *US DOE Heath and Environmental Risk Analysis Program (HERAP)*. Report No. DOE/EV/10731-T1. Contract No. AC02-81EV10731. NTIS no.: PC A08/MF A01. Energy and Environmental Policy Center, John F. Kennedy School of Government, Harvard University, Cambridge, MA, 1983.

Percivall G., (Ed.) *OGC Reference Model, Version 0.1.3*. Open Geospatial Consortium Inc., Wayland, MA, 2003. Available at: http://portal.opengeospatial.org/files/?artifact_id = 3836 (accessed April 2007).

Peters A., Dockery D.W., Muller J.E., and Mittleman M.A. Increased particulate air pollution and the triggering of myocardial infarction. *Circulation* 2001a: 103: 2810–2815.

Peters A., Fröhlich M., Döring A., Immervoll T., Wichmann H.-E., Hutchinson W.L., Pepys M.B., and Koenig W. Particulate air pollution is associated with an acute phase response in men. *Eur Heart J* 2001b: 22: 1198–1204.

Peters J.M., Avol E., Navidi W., London S.J., Gauderman W.J., Lurmann F., Linn W.S., Margolis H., Rappaport E., Gong H., and Thomas D.C. A study of twelve southern California communities -with differing levels and types of air pollution. I. Prevalence of respiratory morbidity. *Am J Respir Crit Care Med* 1999: 159(3): 760–767.

Polman C.H., Reingold S.C., Edan G., Filippi M., Hartung H.P., Kappos L., et al. Diagnositic criteria for multiple sclerosis: 2005 revisions to the "McDonald Criteria". *Ann Neurol* 2005: 58: 840–846.

Pope III C.A. Respiratory disease associated with community air pollution and a steel mill, Utah Valley. *Am J Public Health* 1989: 79: 623–628.

Pope III C.A. Respiratory hospital admissions associated with PM_{10} pollution in Utah, Salt Lake, and Cache Valleys. *Arch Environ Health* 1991: 46: 90–97.

Pope III C.A., Burnett R.T., Thun M.J., Calle E.E., Krewski D., Ito K., and Thurston G.D. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *J Am Med Assoc* 2002: 287: 1132–1141.

Pope III C.A., Dockery D.W., Spengler J.D., and Raizenne M.E. Respiratory health and PM_{10} pollution: a daily time series analysis. *Am Rev Respir Dis* 1991: 144: 668–674.

Pope III C.A. Particulate pollution and health: a review of the Utah Valley experience. *J Expo Anal Environ Epidemiol* 1996: 6: 23–34.

Pope III C.A., Muhlestein J.B., May H.T., Renlund D.G., Anderson J.L., and Horn B.D. Ischemic heart disease events triggered by short-term exposure to fine particulate air pollution. *Circulation* 2006: 114: 2443–2448.

Pope III C.A., Schwartz J., and Ransom M.R. Daily mortality and PM 10 pollution in Utah Valley. *Arch Environ Health* 1992: 47: 211–217.

Pope III C.A., Thun M.J., Namboodiri M.M., Dockery D.W., Evans J.S., Speizer F.E., and Heath J.C.W. Particulate air pollution as a predictor of mortality in a prospective study of US adults. *Am J Respir Crit Care Med* 1995: 151: 669–674.

Pope C.A., Rodermund D.L., and Gee M.M. Mortality effects of a copper smelter strike and reduced ambient sulfate particulate matter air pollution. *Environ Health Perspect* 2007: 115: 679–683.

Poser C.M., Paty D.W., and Scheinber L., et al. New diagnostic criteria for multiple sclerosis: guidelines for research protocols. *Ann Neurol* 1983: 13: 227–231.

Raizenne M., Neas L.M., Damokosh A.I., Dockery D.W., Spengler J.D., Koutrakis P., Ware J.H., and Speizer F.E. Health effects of acid aerosols on North American children: pulmonary function. *Environ Health Perspect* 1996: 104(5): 506–514.

Ritz B., Tager I., and Balmes J. Can lessons from public health disease surveillance be applied to environmental public health tracking? *Environ Health Perspect* 2005: 113: 243–249.

Roberts E.M., English P.B., Wong M., Wolff C., Valdez S., Van den Eeden S.K., *et al.* Progress in pediatric asthma surveillance II: geospatial patterns of asthma in Alameda County, California. *Preventing Chronic Disease* 2006b: 3(3): A92.

Samet J.M., Zeger S.L., Dominici F., Curriero F., Coursac I., Dockery D.W., Schwartz J., and Zanobetti A. The National Morbidity, Mortality, and Air Pollution Study. Part II: morbidity and mortality from air pollution in the United States. *Res Rep Health Eff Inst* 2000a: 94: 5–70.

Samet J.M., Zeger S., Dominici F., et al. *The National Morbidity, Mortality, and Air Pollution Study (NMAPS). Part I. Methods and Methodological Issues*. Health Effects Institute, Cambridge, MA, 2000b (Report no. 94).

Schimmel H., and Greenburg L. A study of the relation to pollution to mortality in New York City, 1963–1968. *J Air Pollut Control Assoc* 1972: 22: 607–616.

Schwartz J. Lung function and chronic exposure to air pollution: a cross-sectional analysis of NHANES II. *Environ Res* 1989: 50(2): 309–321.

Schwartz J. Air pollution and hospital admissions for heart disease in eight US counties. *Epidemiology* 1999: 10: 17–22.

Schwartz J., and Dockery D.W. Particulate air pollution and daily mortality in Steubenville, Ohio. *Am J Epidemiol* 1992: 135: 12–19.

Schwartz J., and Marcus A.H. Mortality and air pollution in London: a time series analysis. *Am J Epidemiol* 1990: 131: 185–194.

Schwartz J. Particulate air pollution and chronic respiratory disease. *Environ Res* 1993: 62(1): 7–13.

Teutsch S. Considerations in planning a surveillance system. In: Teutsch S., and Churchill R. (Eds.). *Principles and Practice of Public Health Surveillance*. Oxford University Press, New York, 2000, pp. 17–29.

Thomas P.T., and Zelikoff J.T. Air pollutants: modulators of pulmonary host resistance against infection. In: Holgate S.T., Samet J.M., Koren H.S. and Maynard R.L. (Eds.). *Air Pollution and Health*. Academic Press, London UK, 1999, pp. 357–379.

Thurston G.D. The use of nationwide registries to assess environmental health effects. *Testimony before the US Senate Committee on Health, Education, Labor, and Pensions*, Subcommittee on Public Health. March 6 Washington, DC, 2002: http://help.senate.gov/Hearings/2002_03_06/2002_03_06.html.

Thurston G.D., Ito K., Kinney P.L., and Lippmann M. A multi-year study of air pollution and respiratory hospital admissions in three New York State metropolitan areas: results for 1988 and 1989 summers. *J Expo Anal Environ Epidemiol* 1992: 2: 429–450.

Tolbert P.E. Invited commentary: heterogeneity of particulate matter health risks. *Am J Epidemiol* 2007: 166: 889–891.

Tonne C., Beevers S., Armstrong B.G., Kelly F., and Wikkinson P. Air pollution and mortality benefits of the London Congestion Charge: spatial and socioeconomic inequalities. *Occup Environ Med* 2008. [Advance online publication, 28 February 2008. doi:10.1136/oem.2007036533]

Uitdehaag B.M., Kappos L., Bauer L., Freedman M.S., Miller D., Sandbrink R., and Polman C.H. Discrepancies in the interpretation of clinical symptoms and signs in the diagnosis of multiple sclerosis. A proposal for standardization. *Mult Scler* 2005: 11: 227–231.

Ungar W., Chapman K., and Santos M. Assessment of a Medication-Based Asthma Index for Population Research. *Am J Respir Crit Care Med* 2002: 165: 190–194.

US Senate Bill S. *1442: Coordinated Environmental Health Network Act of 2005* 2005 (http//www.help.senate.gov/Bills/Bills_2006_109_b.htm).

Vella V. A review of the etiology of multiple sclerosis. *Ital J Neurol Sci* 1984: 5: 347–356.

Ware J.H., Thibodeau L.A., Speizer F.E., Colome S., and Ferris Jr B.G. Assessment of the health effects of atmospheric sulfur oxide and particulate matter: evidence from observational studies. *Environ Health Perspect* 1981: 41: 255–276.

Willer C.J., and Ebers G.C. Susceptibility to multiple sclerosis: interplay between genes and environment. *Curr Opin Neurol* 2000: 13: 241–247.

Woodruff T.J., Grillo J., and Schoendorf K.C. The relationship between selected causes of postneonatal infant mortality and particulate air pollution in the United States. *Environ Health Perspect* 1997: 105(6): 608–612.

Woodruff T.J., Parker J.D., and Schoendorf K.C. Fine particulate matter ($PM_{2.5}$) air pollution and selected causes of postneonatal infant mortality in California. *Environ Health Perspect* 2006: 114(5): 786–790.

Zanobetti A., and Schwartz J. The effect of particulate air pollution on emergency admissions for myocardial infarction: a multicity case-crossover analysis. *Environ Health Perspect* 2005: 113(8): 978–982.

Zelikoff J.T., Chen L.C., Cohen M.D., Fang K., Gordon T., Li Y., Nadziejko C., and Schlesinger R.B. Effects of inhaled ambient particulate matter on pulmonary antimicrobial immune defense. *Inhal Toxicol* 2003: 15(2): 131–150.