Published in final edited form as:
Environ Int. 2017 March ; 100: 62–78. doi:10.1016/j.envint.2016.12.019.

Accountability studies of air pollution and health effects: lessons learned and recommendations for future natural experiment opportunities

David Q. Rich, ScD
Departments of Public Health Sciences and Environmental Medicine, University of Rochester School of Medicine and Dentistry, Rochester, New York

Abstract

To address limitations of observational epidemiology studies of air pollution and health effects, including residual confounding by temporal and spatial factors, several studies have taken advantage of 'natural experiments', where an environmental policy or air quality intervention has resulted in reductions in ambient air pollution concentrations. Researchers have examined whether the population impacted by these air quality improvements, also experienced improvements in various health indices (e.g. reduced morbidity/mortality). In this paper, I review key accountability studies done previously and new studies done over the past several years in Beijing, Atlanta, London, Ireland, and other locations, describing study design and analysis strengths and limitations of each. As new 'natural experiment' opportunities arise, several lessons learned from these studies should be applied when planning a new accountability study. Comparison of health outcomes during the intervention to both before and after the intervention in the population of interest, as well as use of a control population to assess whether any temporal changes in the population of interest were also seen in populations not impacted by air quality improvements, should aid in minimizing residual confounding by these long term time trends. Use of either detailed health records for a population, or prospectively collected data on relevant mechanistic biomarkers coupled with such morbidity/mortality data may provide a more thorough assessment of if the intervention beneficially impacted the health of the community, and if so by what mechanism(s). Further, prospective measurement of a large suite of air pollutants may allow a more thorough understanding of what pollutant source(s) is/are responsible for any health benefit observed. The importance of using multiple statistical analysis methods in each paper and the difference in how the timing of the air pollution/outcome association may impact which of these design features is most important is also discussed. Based on these and other lessons learned, researchers may provide a more epidemiologically rigorous evaluation of cause-specific health impacts of an air quality intervention or action.

Corresponding Author: David Q. Rich, ScD, University of Rochester School of Medicine and Dentistry, Department of Public Health Sciences, 265 Crittenden Boulevard, CU 420644, Rochester, NY 14642, Phone: 585-276-4119, Fax: 585-424-1469, david_rich@urmc.rochester.edu.

Keywords

Air pollution; epidemiology; accountability; natural experiment; confounding

INTRODUCTION

Numerous studies have reported associations between air pollution and cardiorespiratory mortality and morbidity,[1-5] primarily focusing on respiratory outcomes (e.g. asthma exacerbation, asthma incidence, wheezing), cardiovascular outcomes (e.g. myocardial infarction, stroke, arrhythmia, heart failure), and lung cancer. More recently they have also studied reproductive (e.g. preterm birth, fetal growth restriction, stillbirth, pregnancy complications) and neurologic outcomes (e.g. autism). Others have studied air pollution and adverse changes in relevant biomarkers of physiological and biochemical function (e.g., heart rate variability, lung function, inflammation, oxidative stress).[1-3] However, criticisms of these studies have included exposure misclassification and its resulting bias towards the null, and residual confounding by both temporal and spatial factors including time trends in exposure and health (e.g. hour of the day, weekday, season, long term time trends), socioeconomic status and access to health care, smoking, other neighborhood factors, exposures to other toxicants, and occupational exposure to pollution. A central limitation of these epidemiology studies is that they are observational in nature, and the investigator cannot control who is exposed and who is not, and therefore can only measure many of these potential confounders and try to control for them in his/her analytic models.

However, other studies have taken advantage of ‘natural experiments’ where a large scale air pollution reduction or composition change was observed in a community, region, or country following a legislative mandate, during a large scale sporting event, during an employee strike at a large industrial facility, etc. Investigators then examined whether improvements in morbidity/mortality were observed concurrent to these periods of lower air pollution levels. These natural experiments come as close to a laboratory-controlled experiment as possible in epidemiology/observational science, and allow one to observe whether the same pollutant/health endpoint associations observed in the studies described above (increased pollution concentrations associated with increased morbidity and/or mortality) exist when pollution is reduced (i.e. decreased pollution concentrations associated with decreased morbidity and/or mortality).

These “accountability” studies evaluate the extent to which an air pollution improvement or regulation in a city or region beneficially impacted human health. Most were opportunistic in nature and used already existing health and pollution data to evaluate whether an action resulted in air quality and/or health improvements in impacted populations. In 2003, the Health Effects Institute published a monograph providing a conceptual framework defining the ‘chain of accountability’ and outlining how such studies can be done (Figure 1).[6] For example, that chain describes how a regulatory action can impact emissions, ambient air quality, exposure/dose, and then human health. Accountability studies can ask a series of questions based on this chain. For example, has a regulatory action or controls reduced emissions? Have pollutant concentrations been reduced due to source controls and emissions reductions? Has this led to reduced subject exposures? Have these reduced subject exposures led to reduced doses? Have health risks declined as a result of any of these actions on previous components of the chain? In 2010, HEI hosted a workshop to discuss the current state of accountability research conducted by investigators funded by HEI, with a summary of lessons learned described by van Erp et al 2012.[7] Currie et al (2011) also provided a review of 10 such studies assessing whether such air pollution reductions impacted infant's and children's health.[8] This review provides an update, describes both older and more recent studies, and provides a set of epidemiologic recommendations for future investigators that want to conduct such studies when a new natural experiment opportunity arises. Studies were identified using www.pubmed.org and search terms such as natural experiment, accountability, PM, air pollution and health, etc. Others studies were identified if they were discussed in one of the papers found via this PubMed search.

Although studies have assessed whether dramatic increases in air pollutant concentrations due to accidents/disasters (e.g. World Trade Center attacks in New York City, oil tank explosion in Gulen, Norway) have impacted health,[9,10] these are not included in this review. The focus of this paper is a discussion of those natural experiments where an air pollution *reduction* was observed or expected, and whether that reduction resulted in a health benefit was examined. Below I summarize accountability studies assessing health effects of 1) industrial pollution sources and large industrial shutdowns, 2) residential heating fuel and changes in fuel type, 3) traffic pollution and changes in motor vehicle fuel composition or plans to reduce traffic congestion, 4) national or state wide air pollution reductions, and 5) reductions in air pollutant levels during large scale sporting events. Listed in Table 1 and discussed below for each study, are descriptions of the study location, the natural experiment scenario, the 'chain of accountability' question assessed, study objective(s), the study's main findings, and its key strengths and limitations. The strengths and limitations (i.e. sources of bias and residual confounding) listed in Table 1 may not be a complete list for each study. There may be other sources of residual confounding in these observational studies, even though each study has made attempts to identify them, minimize their effects, and/or discuss potential impacts on their effect estimates. Last, a discussion of key problems/strengths across these studies and recommendations for the designs of future accountability studies, if such opportunities arise, is provided.

INDUSTRIAL POLLUTION SOURCES AND INDUSTRIAL SHUTDOWNS

Utah Valley Steel Mill

One of the first such natural experiments used to assess air pollution/health associations occurred when a large steel mill closed in the Utah Valley from August, 1986 to September 1987 (13 months) in response to an employee strike. Until that time, the largest contributor to PM_{10} (particles <10 μm in aerodynamic diameter) concentrations in the valley was that large steel mill, with multiple occasions when the 24 hour PM_{10} concentration was >150 μg/m^3 during this pre-shutdown period. However, when the mill shut down during the employee strike, average PM_{10} concentrations were ~43% lower (from 90 μg/m^3 to 51 μg/m^3), and there were no 24 hour averages >150 μg/m^3.[11] Compared to the periods before and after the strike when pollution levels were higher, Pope and colleagues reported reductions in hospital admissions for pneumonia, pleurisy, bronchitis, and asthma,[11] childhood school absenteeism,[12] bronchitis and asthma admissions for pre-school aged children,[13] respiratory symptoms and reduced peak expiratory flow,[14] and total non-accidental mortality, respiratory mortality, and cardiovascular mortality during the period of reduced air pollution levels (i.e. during the strike).[15] For these analyses, Pope and colleagues used hospital admissions data and daily air pollution measurements from an air pollution monitoring station in Lindon, Utah. Statistical analyses were generally time-series analyses utilizing Poisson regression models, adjusting for daily temperature and relative humidity.

Strengths and limitations—A strength of these analyses was the study design allowing for an A-B-A comparison where the mortality rate during the air pollution reduction period (i.e. strike and mill shutdown; Pollution Level B) was compared to the mortality rates both before (Pollution Level A) and after the shutdown (Pollution Level A) when the air pollution levels were much higher. This allowed for a more robust control of confounding by long term time trends in pollution and mortality then just an A-B analysis (i.e. comparison of mortality rate during shutdown when air pollution levels are reduced [Pollution Level B] only to the mortality rate before shutdown when air pollution levels were higher [Pollution Level A]).

Utah Valley Steel Mill – Birth Outcomes

Taking advantage of this same natural experiment, Parker et al. (2008) examined whether there was an improvement in the rate of preterm birth (<37 weeks gestational age) during this 13 month period of reduced air pollutant concentrations.[16] They used birth records from 1984 to 1990 from the National Center for Health Statistics for pregnancies where the mother resided in either Utah County (Utah Valley including Provo and Orem) or other Utah counties, including only singleton births of non-Hispanic white, married women with complete data on birth weight and gestational age on the birth certificate. Parker et al (2008) categorized all Utah Valley pregnancies as having exposure during the first trimester, second trimester, third trimester, first and second trimesters, second and third trimesters, and whole pregnancy (all 3 trimesters) during the 13 month strike. They then compared the rate of preterm birth in each of these exposure windows to pregnancies of Utah Valley women where the comparable portion of the pregnancy was during the same calendar dates in the 2 years before or the 2 years after the strike. Second, they repeated this analysis for births/ pregnancies to women residing in other Utah counties during the same time. They found that Utah Valley mothers who were already pregnant at the time of mill closure were less likely to deliver prematurely than mothers pregnant before or after the closure. They saw no such pattern in non-Utah Valley births. Further, the largest effects were observed for pregnancies with their 2nd trimester during the mill closure (RR=0.86; 95% CI =0.75, 0.98). However, they did not find a significant reduction in birth weight during the steel mill closure.[16]

Strengths and Limitations—Strengths of the analysis included the A-B-A design and its control of time trends in air pollution and preterm birth. Further, the simultaneous assessment of these same temporal comparisons in other Utah counties provided control for background changes in the preterm birth rate occurring during this time in Utah independent of any air pollution change. However, the study also had limitations. First, the time period during which the air pollution reduction occurred in the valley was long relative to the length of pregnancy (~13 months). This made examination of specific windows of pregnancy (e.g. 1st month or last month), when reduced pollutant concentrations might result in improved birth outcomes, difficult. Identifying specific windows of pregnancy, when air pollution impacts fetal growth, may help identify potential underlying physiologic mechanisms. Second, there was no examination of pollution levels in the other non-Utah Valley counties. Last, there was no complementary direct assessment of whether air pollution levels in Utah Valley during these years were significant predictors of preterm birth in these Utah Valley residents to provide evidence that changes in air pollution were driving these preterm birth rate changes.

Four Southwestern States – Copper Smelters

Using a similar study design to the original Utah Valley analyses, Pope et al (2007) also examined whether a nationwide copper smelter strike from July 1967 to April 1968 and its reduction in air pollution (e.g. ~60% reduction in suspended sulfate particles), resulted in a decreased mortality rate in four southwestern US states (New Mexico, Arizona, Utah, and Nevada). Before the strike, Pope et al. (2007) estimated that ~90% of sulfate emissions in these four states were made by copper smelters, and that a major component of $PM_{2.5}$ (particulate matter <2.5 um in aerodynamic diameter) concentrations was sulfate, with an ~2.5 μg/m^3 reduction in suspended sulfate concentration observed regionally during the strike.[17] Using a similar Poisson regression modeling approach as in the Utah Valley study, they regressed monthly mortality counts against indicator variables for the strike period, adjusting for time trends, mortality counts in bordering states, and nationwide mortality counts for influenza/pneumonia, cardiovascular disease, and other respiratory mortality (as a means of controlling for both national and regional trends in cardiorespiratory mortality). They reported a 2.5% decrease in mortality (95% CI = 1.1%, 4.0%) associated with the copper smelter strike in these 4 states, again suggesting a health benefit in the population presumably driven by a reduction in air pollution emitted by copper smelters in these states.[17]

Strengths and Limitations—This analysis had similar strengths to those listed above for the Utah Valley Steel Mill analyses. It also adjusted for regional and national background cardiovascular and respiratory (both influenza/pneumonia and other non-infectious respiratory diseases) mortality rates, which should have further reduced residual confounding by changes in the underlying disease rate occurring across the whole region/nation.

RESIDENTIAL HEATING FUEL AND CHANGE IN HEATING FUEL TYPE

Ireland 1 - Dublin

In 2002, Clancy and colleagues described their assessment of the natural experiment in Dublin, Ireland centered on a regional ban on the marketing and sale of coal for residential combustion in 1990.[18] Coal, primarily bituminous coal, was used increasingly in the 1980's in Ireland, instead of oil, for residential space and water heating, resulting in adverse impacts on air quality,[19] with some suggesting concomitant increases in adverse health effects including in-hospital respiratory mortality.[20] Thus, the Irish government banned the marketing, sale, and distribution of coal in the city of Dublin on September 1, 1990. Using national mortality records and measurements of air pollutants (black smoke and SO_2) from 6 residential monitoring stations in Dublin, Clancy et al (2002) compared both air pollution levels and cause-specific mortality rates in the 6 years before and after the ban. They reported improvements in the concentrations of black smoke in Dublin (−71%; 50.2 μg/m^3 to 14.6 μg/m^3 across all seasons) with the largest changes observed in the winter (~75% reduction; 85.4 μg/m^3 to 21.4 μg/m^3), when use of coal for heating was greatest. Concurrent to these air pollution changes, they also reported improvements in age-standardized, cause-specific mortality rates (total non-trauma: −6% [95% CI = −4%, −7%], respiratory: −16% [95% CI = 12%, 19%], and cardiovascular: −10% [95% CI =−8%, −13%]), after adjustment for temperature, relative humidity, influenza epidemics, and cause-specific mortality rates from the rest of Ireland.[18]

Strengths and Limitations—The strengths of this design included the use of readily available, nationally-maintained mortality data and air pollution measurements from several monitoring locations in Dublin, without the need for more costly and time consuming prospective data collection. Further the study provided assessment of changes in cause-specific mortality after adjustment for influenza epidemics, thereby reducing potential confounding by other causes of respiratory mortality. Similar to the Utah Valley studies, Clancy et al (2002) also attempted to control for background mortality rates across Ireland. However, some limitations in this approach and data used to adjust for background rates are noted and discussed below.

Ireland 2 - Cork, Five Smaller Cities, and Dublin re-analysis

The Irish government legislated similar coal sale bans in other cities/counties of Ireland in 1995 (Cork County Borough) and five smaller cities in 1998 (Arklow, Drogheda, Dundalk, Limerick, and Wexford), providing another natural experiment opportunity to be used to replicate the Dublin findings. As part of this study, Goodman et al (2009) reported a 49% reduction in black smoke concentrations in Cork (from 33.7 μg/m^3 to 17.2 μg/m^3), and 48% to 61% reductions in the 5 smaller cities of the 1998 ban, when comparing the 5 years after the ban to the 5 years before each ban.[21] Dockery et al, (2013) then examined whether changes in cause-specific mortality rates (using similar national mortality records) and cause-specific morbidity rates (using hospital admissions data for a set of hospitals in Cork County Borough) were observed. As part of the mortality analyses, they also re-examined whether the Dublin ban was associated with reductions in cause-specific mortality rates, using alternative statistical methods and confounder control, as well as an extended study period (here using 10 years before and after the ban). When using interrupted time-series analyses with this longer study period, and use of cause-specific mortality rates from counties (i.e. Coastal Counties of Clare, Donegal, Kerry, Leitrim, Mayo, and Sligo) thought to be unaffected by the Coal Sale bans in Dublin (1990), Cork (1995), and other cities (1998), they reported no reduction in total mortality (−1% [95% CI = −6%, 4%]; −4% [95% CI = −10%, 1%]; 0% [95% CI = −3%, 4%] respectively), or cardiovascular mortality (0% [95% CI = −9%, 10%]; −4% [95% CI = −12%, 6%]; −1% [95% CI = −6%, 4%] respectively) for these three bans. However, they did still find significant reductions in respiratory mortality following the 1990 coal sale ban in Dublin (−17%; 95% CI = −24%, −8%). Further, they reported non-significant reductions following the Cork 1995 coal sale ban (−9%; 95% CI = −18%, 1%), and the other cities in 1998 (−3%; 95% CI = −8%, 3%).[22] As a comparison, they estimated the change in total, cardiovascular, and respiratory mortality in 12 “Midland” counties (Cavan, Monaghan, Meath, Laois, Longford, Offaly, Westmeath, North and South Tipperary, Carlow, Kilkenny, and Roscommon), which were presumed to be unaffected by the coal sale ban, associated with each of the bans. Adjusting for the same covariates as in Dublin, including the cause-specific mortality rates in the coastal counties, the 1990 ban in Dublin was associated with small non-significant decreases in total (−3%; 95% CI=−8%, 3%), cardiovascular (-2%; 95% CI = −10%, 7%), and respiratory mortality (−2%; 95% CI = −12%, 8%). There were also similar small non-significant changes in cause-specific mortality rates associated with the 1995 Cork and 1998 5 cities coal sale bans.[22]

Strengths and Limitations—In this revised analysis of the Dublin coal sale ban,[22] the addition of a control population of other counties to assess whether the 1990 Dublin ban had mortality effects outside Dublin, and the inclusion of cause-specific mortality rates in the coastal counties provided likely better control of confounding by long term time trends and background mortality rates than in the original analysis,[18] and are a strength of the analysis. However, these revised Dublin findings with regard to total and cardiorespiratory mortality were in direct contrast to those reported by Clancy et al (2002) in the first analysis, and highlight the difficulty in retrospectively adjusting for background trends in mortality rates that are occurring at the same time as the reductions in air pollution. Further, they highlight the importance of having a control population with comparable health experiences during the study period that was also not affected by the intervention and its reduced air pollutant concentrations to assess if any residual confounding remains. As discussed by Cox and Popken (2015),[23] the lack of an adequate control group that experienced the same trends in mortality (here a decline in cardiovascular and total mortality over the past several years), but not the improvements in air quality, makes any conclusion about the impact of an intervention on mortality (here total or cardiovascular mortality) less convincing.

This Ireland natural experiment study also attempted to assess whether morbidity rates were impacted by the coal sale ban in Cork (via hospital admissions records). However, the data system used to compile and store the hospital admissions records began use at essentially the same time as the coal sale ban. Thus the first few years appeared to have lower number of admissions compared to later years, making it difficult to estimate a change in cause specific hospital admission rates associated with the coal sale ban. Further, there was not a control population on which to run similar analyses as described above for mortality.[22]

NATIONAL OR STATE WIDE AIR POLLUTION REDUCTIONS

German reunification

Following the reunification of Germany in 1990, concentrations of total suspended particulates and sulfur dioxide concentrations decreased substantially.[24] Prior work in Erfurt, Germany (formerly in East Germany) reported that accumulation mode particle

(100-500nm) concentrations decreased, but nucleation mode particles (10-30nm) increased since 1990 in this region of Germany.[25] They argued that these concentration changes (mostly reductions) occurring since reunification, may be due to reduced emissions from stationary sources, increases in motor vehicle use and emissions, and due to air quality regulations targeted at the reduction of larger particles.[25-27] Heinrich et al (2002) then conducted three cross-sectional surveys (1992-1993; 1995-1996, and 1998-1999) of children (5-14 years of age) living in 3 areas in the former East Germany, and used a 2-stage modeling approach to estimate the change in several respiratory disorders (bronchitis, otitis media, sinusitis, frequent colds, infections, cough, and shortness of breath) associated with changes in TSP and SO_2 concentrations during this time. Across these time periods, the prevalence of each of these respiratory conditions declined. They reported that each 50 $\mu g/m^3$ increase in TSP was associated with a significant 202% increase in the prevalence odds of bronchitis (95% CI = 72%, 429%), 158% increase in the prevalence odds of sinusitis (95% CI = 0%, 565%), and 90% increase in the prevalence odds of frequent colds (95% CI= 17%, 209%), after adjustment for age, gender, parental education as a marker of socioeconomic status, parental atopy, home dampness or molds, gas cooking, second hand smoke exposure, and cat contact. They reported similar sized effect estimates associated with each 100 $\mu g/m^3$ increase in SO_2 concentrations. Thus, they concluded that these reductions in the prevalence of common respiratory illnesses in children following the reunification of Germany, were, in part, driven by the substantial declines in air pollution levels observed.[28]

Strengths and Limitations—The strengths of this study include the measurements of TSP, sulfur dioxide, as well as particle number counts of size fractioned particles (here accumulation mode particles and nucleation mode particles) allowing a greater understanding of pollution sources and how they changed following reunification. Further, the study contacted schools and daycare centers to recruit children that had lived in the study area for at least 2 years and had not relocated to the current house from than 2km away, and attempted to re-contact each child for the 2 successive surveys. This assessment of some of the same children across the three time periods, and the prospective collection of data are strengths of the study. As noted by the study authors, however, the changes to a western life-style (e.g. diet, medication use, access to care) may also have impacted the prevalence of these respiratory outcomes in children across the 3 time periods. Thus, residual confounding by temporal changes in these factors may also explain some of the reductions in respiratory illness prevalence. Further, without a control population as in the Ireland analysis described above,[22] residual confounding by long term time trends is possible.

Germany – power plant scrubber installations

Using birth and death records maintained by a state agency, and SO_2 data from 553 monitoring station to estimate county wide SO_2 concentration averages, Luechinger (2014) examined whether there were reductions in infant mortality in Germany from 1985 to 2003 following mandated installations of scrubbers at power plants. Using fixed effects regression models adjusted for TSP pollution, rural and urban trends, reunification effects, weather, and demographic development, every 1 $\mu g/m^3$ decrease in annual, county average SO_2

concentration was associated with a reduction in infant mortality of 2.6 lives per 100,000 births.[29]

Strengths and Limitation—Although, this study used a large national dataset and adjusted for multiple potential temporal and spatial confounders, they assessed only SO_2, and no other pollutants (e.g. $PM_{2.5}$, sulfate). Since this was a national evaluation, there was no control population possible to adjust for background infant mortality rates from a similar area that was not impacted by this power plant scrubber installation. Thus, residual confounding by changes in diet, other environmental conditions, cultural changes, as well as long term time/health trends could remain.[29]

California air quality policies

Air quality-control policies in California, targeted at mobile and stationary sources and including fuel and consumer-product reformulations, have resulted in decreased air pollutant concentrations over the past few decades. Gauderman et al (2015) used children's lung function measurements from 3 cohorts of the Children's Health Study, and examined whether lung-function development from 11 to 15 years of age (operationalized as changes in FEV_1 and FVC over that 4 year period) was associated with decreases in concentrations of several air pollutants (NO_2, $PM_{2.5}$, PM_{10}, and O_3) during those same 4 years of age.[30] Decreases in the concentrations of NO_2 (14.1 ppb), $PM_{2.5}$ (12.6 μg/m^3), and PM_{10} (8.7 μg/m^3) across the study's 5 California communities were associated with 91.4 ml (95% CI = 47.9, 134.9), 65.5 ml (95% CI = 17.1, 113.8), and 65.5 ml (95% CI = 27.2, 103.7) increases in FEV_1 from 11 to 15 years of age, respectively, and 168.9 (95% CI = 127.0, 210.7), 126.9 ml (95% CI = 65.7, 188.1), and 113.0 ml (95% CI = 60.0, 166.1) improvements in FVC from 11 to 15 years of age, respectively, after adjustment for sex, race, Hispanic ethnicity, height, body mass index, and respiratory tract illness on the day of assessment. Improvements in 4 year growth of FEV_1 and FVC were observed in both boys and girls, and in children with and without asthma.[30]

Strengths and limitations—Strengths of the study include use of a large, well characterized cohort of children with prospective data collection, and available air quality data for the entire study period. Dockery and Ware (2015) argue that this study provides corroborating evidence of air pollution impacts on pulmonary function since these analyses are based on comparisons within communities, and thus are not confounded by differences in characteristics between communities. However, they argue that changes in racial/ethnic composition and other characteristics of the cohort over the study period (e.g. use of healthcare relative to health insurance copayments) may result in residual confounding.[31] Similar to other studies discussed above, it was not possible to use another community with a similar health history, but dissimilar air pollutant concentration patterns over time (a control population), as a means to control for this residual confounding.[30]

TRAFFIC POLLUTION, CHANGES IN VEHICLE FUEL COMPOSITION, TRAFFIC CONGESTION REDUCTION PLANS

Hong Kong

Hedley and colleagues assessed whether a reduction in the sulfur content of fuel oil used in power plants and road vehicles in Hong Kong from 1990 to 1995 was associated with reduced mortality among Hong Kong residents. Compared to before the intervention, concentrations of SO_2 fell by ~80% (Kwai Tsing [highly polluted district]: peak of 88-101 $\mu g/m^3$ in 1989 and 113-136 $\mu g/m^3$ in 1990, to 23-26 $\mu g/m^3$ in 1991) and sulfate by 38% after the intervention (Kwai Tsing: 12.5 $\mu g/m^3$ in 1989 to 7.7 $\mu g/m^3$ in 1991). There were no such large changes in other pollutant concentrations.[32] Hedley et al (2002) used census data (ICD-9 codes) to provide monthly counts of all cause, respiratory, cardiovascular, and other causes of mortality, district board resident populations to estimate the population covered by the 5 monitoring stations, and time series methods (Poisson regression) to estimate the change in seasonal mortality after the intervention. They reported small but significant reductions in seasonal counts of respiratory (−3.9%), cardiovascular (−2.0%), and total mortality (−2.1%), compared with predicted seasonal values, with the largest declines in the first 12 months after the invention.[33] This is consistent with their previous work showing greater reductions in bronchial hyper-responsiveness and chronic bronchitis symptoms, after the intervention, in children living in a polluted district compared to children living in an unpolluted district.[32,34]

Strengths and Limitations—Although this study examined whether the seasonal pattern in mortality changed after this fuel intervention was implemented, it did not control for influenza and other infectious disease epidemics, which may contribute to seasonal mortality, and did not control for mortality rates in other similar populations (if possible) not affected by this fuel intervention. However, given the tropical location of Hong Kong, residual confounding by influenza may not have a large impact.[33]

London, England

Starting in early 2003, the mayor of London, England implemented a congestion charging zone in central London, which charged drivers a small fee (~8 pounds in 2008) per day to drive into the zone during working hours. As of 2008, congestion had improved with a 26% reduction in the excess delay per kilometer, and a reduced traffic volume during the charging hours, compared to before the charging scheme was implemented.[35] Kelly et al (2011) then examined whether the London low emission zone had beneficial impacts on air pollution levels, traffic patterns, and traffic volume in central London. Their modeling efforts predicted small to moderate reductions in NO_x (−3.8% by 2008 and −7.3% by 2012) and PM_{10} emissions (−2.6% by 2008 and −6.6% by 2012).[36]

Another goal of this project was to assess the suitability of using a large primary care database of London residents to assess whether this congestion charging scheme resulted in any health benefits. Pilot analyses suggested essentially no link between NOx exposures and the prevalence or incidence of cardiorespiratory outcomes (asthma, chronic obstructive pulmonary disease, wheeze, hay fever, upper and lower respiratory tract infections, ischemic heart disease, heart failure, and atrial fibrillation). Further, using a quantitative risk assessment approach (i.e. not measuring health outcomes directly), they estimated the mortality benefit of this reduction in the low emission zone of London, to be 183 years of life gained per 100,000 people.[37]

Strengths and Limitations—Kelly et al (2011) documented some difficulties in using primary care medical records for a health analysis, as some primary care providers were concerned about losses in confidentiality due to geographic patterns of NO_2 and PM_{10} concentrations relative to residences of their patients in the dataset. Therefore, they restricted analyses to NOx only in their pilot analyses. However, in further power calculations, they determined that these health data may be valuable for future analyses, as they would provide the power to detect a 5% decline in various health measures (e.g. medication prescriptions, respiratory consultations).[38] Kelly et al (2011) also did not compare these traffic reductions or changes in cardiorespiratory outcomes (from 2003 to 2008) to another London area for the same time period. This would have provided further evidence that the observed traffic changes, and thus estimated air pollution reductions, were a result specifically of the congestion charging scheme.

New Jersey and Pennsylvania

From 1997 to 2000, toll plazas on highways in New Jersey and Pennsylvania installed the E-ZPass toll collection system. E-ZPass allows drivers to pass through the toll without stopping while electronically collecting the toll payment for each car, thereby reducing traffic congestion, idling, and likely air pollutant concentrations at and near the toll plazas. Using vital statistics data to provide information on births (i.e. gestational age, birth weight, etc.) to mothers living near these 98 toll plazas for the 3 years before and after E-ZPass installation, Currie and Walker (2011) assessed whether there were reduced numbers of preterm births (<37 weeks gestational age) and low birth weight babies (<2500 g) after E-ZPass was installed, among mothers living near (within 2 km) and farther away (2-10 km) from a toll plaza. Using difference-in-differences models, the introduction to E-ZPass was associated with 11.8% and 10.8% reductions in low birth weight babies and preterm births, respectively, among mothers living within 2 km of the toll plaza relative to mothers living 2-10 km from the toll plazas. Further, they found no difference in the demographic characteristics or housing prices after E-ZPass installation.[39]

Strengths and Limitations—The study had several strengths, including the use of large vital statistics datasets to provide all the health data needed, and the use of a control population to evaluate whether similar changes in birth outcomes were observed in both the populations living near the toll plazas, and the populations living farther away from the toll plazas, thereby minimizing confounding by time trends in birth outcomes and air pollution. Although air pollution monitoring was not done at the sites or the communities within 2 km and 2-10 km from the toll plazas, the assumption is that reductions in traffic congestion and idling times at the toll plazas, and thus reductions in air pollution levels in the areas around the toll plazas may explain these health improvements.

Puget Sound, Washington

In the mid 2000's, Washington State implemented an emissions reduction program to retrofit school diesel school buses with aggressive pollution control technologies to reduce diesel emissions and pollution concentrations on the bus. Adar et al (2015) characterized pollutant exposures and measures of respiratory biomarkers in 275 school bus riders before, during, and after retrofit program in Washington State (2005-2009), and found that fine and ultrafine particle concentrations were reduced in buses using these control technologies by 10-50%.[40] Beatty and Shimshack (2011), however, hypothesized that since school buses travel through residential areas in the school district, perhaps ambient air pollutant concentrations would be reduced in the community. Using state hospital admissions data for adults with chronic health conditions and children living in school districts in the Puget Sound area of Washington state, and difference-in-differences models, they found that child residents of the 34 school districts adopting the retrofit program experienced 23% and 37% fewer admissions for bronchitis/asthma and pleurisy/pneumonia, respectively, relative to residents living in non-adopter school districts. [41]

Strengths and Limitations—Similar to Currie and Walker (2011), strengths of the study included the use of a large statewide hospital admissions dataset providing close to complete coverage of all hospital admissions to residents of the Puget Sound school districts/communities under study, and the use of a control population (here residents of the 9 non-adopting school districts), which likely minimized residual confounding by time trends. Further, this study had a similar limitation to Currie and Walker (2011) in that it did not provide estimates of the change in air pollution levels after the retrofit program in both the adopting and non-adopting school districts/communities to provide evidence that the air pollution change was driving this response.

AIR POLLUTION REDUCTIONS AND LARGE SPORTING EVENTS

1996 Atlanta Summer Olympics – Study 1

During the 1996 Atlanta Olympics, a system was implemented to reduce vehicular traffic and congestion so that spectators could reach Olympic events with little delay and to reduce ozone pollution (frequent ozone violations occurred regularly in the summertime in Atlanta). This included a new public transportation system (24 hour a day) including buses as part of a park and ride service, local businesses increasing use of alternative employee work hours (times other than 9-5pm) and telecommuting, closing the downtown area to private automobiles, and using public warning systems to notify people of traffic and air quality problems. This provided an opportunity to examine if reductions in traffic, ozone, and traffic pollution would result in fewer acute care visits and hospitalizations among Atlanta residents. Using datasets of hospitalizations, emergency department visits, and urgent care center visits for residents of 5 metropolitan Atlanta counties aged 1-16 years, Friedman et al (2001) reported reductions in acute care and hospitalizations for asthma, reduced PM_{10}, CO, and O_3 concentrations, and 22.5% reductions in peak morning traffic on week days, during the Olympics compared to the 4 weeks before and 4 weeks after the Games.[42]

Strengths and Limitations—Although the study utilized four datasets for a 5 county area, it did so for only a 73 day study period, resulting in limited statistical power. Further, residual confounding by seasonal patterns in pollution and hospital admissions, residual confounding by Atlanta residents changing their behaviors during the Games (resulting in a reduction in asthma visits and hospitalizations during the Olympics independent of any reduced air pollution effect on asthma), could explain these findings.

1996 Atlanta Summer Olympics – Study 2

Therefore, Peel et al (2010) used a larger dataset of emergency department data from 1995 to 2004 (longer time period to allow better control of season and long term time trends), more spatially diverse air pollution measurements across the metropolitan Atlanta area, and traffic data from 18 sites within the five counties of the Atlanta metropolitan area to reanalyze the health effects of this Olympic traffic reduction campaign. Using Poisson generalized linear regression models, they conducted several analyses contrasting ED visits during the Olympic Games to both the weeks before and after the games in 1996, as well as to ED visits in the years before and after the 1996 Olympic Games. They reported large reductions in ambient ozone concentrations (20%-30%) in Atlanta, but smaller reductions in CO. However, they reported that similar declines in ozone concentration were observed throughout Georgia and the southeastern United States during this time. They also found 2% to 20% reductions in weekday peak morning traffic counts, consistent with, but smaller in magnitude[43] than those of Friedman et al (2001). However, there was no evidence of reduced cardiovascular or respiratory ED visits during the Olympic Games, which was inconsistent with the earlier work by Friedman and colleagues.

Strengths and Limitations—Peel and colleagues utilized a large, well characterized ED dataset in a large urban 5 county population, did so for a longer study period than Friedman et al (2001), and in sensitivity analyses examined whether similar health responses were observed between study subjects residing inside and outside a perimeter highway. However, there were also several limitations. First, the Olympic period was only 17 days, and thus the daily number of ED visits may still not have been large enough to provide ample statistical power. Further, the air pollutant reductions in Atlanta during the Olympic Games may have been too subtle or for too short a time to induce a measureable health effect in the population. Last, although Peel et al (2010) conducted several analyses to control for temporal confounding, assessed whether similar air pollution and traffic reductions were observed in other areas outside of Atlanta, there was a lack of a ‘control’ area/county/city in their health effects analyses with a similar population where such a traffic reduction system was not implemented to control for temporal and seasonal changes in CV and respiratory ED visits. Thus residual confounding could still remain.

2008 Beijing Summer Olympics – Cardiorespiratory outcomes

As a condition for hosting the 2008 Summer Olympic Games, the Chinese government agreed to make Beijing’s air quality (e.g. Pre-Olympic 24-hour mean $PM_{2.5}$ = 100.9 μg/m^3)[44] comparable to that of previous host cities. Thus, a series of aggressive control measures were implemented to achieve this goal, including coal and leaded gasoline bans, tightening of motor vehicle emission standards, relocation and closing or reducing the polluting capacity of industrial facilities, suspension of construction activities, and removal of approximately one-half of the cars (1.5 million cars) from Beijing roads on alternate days.[45] Concentrations of several PM and gaseous co-pollutants differed dramatically before, during, and after the Olympics.[44,46-49] For all pollutants but ozone, concentrations declined substantially from the Pre- to the During-Olympic period (−60% to −18%), and increased substantially from the During-Olympic to the Post-Olympic period (+21% to +197%).[44,46]

Zhang and colleagues conducted a prospective panel study in healthy young medical residents living on or near a hospital campus to examine whether these large drastic air pollutant concentrations declines from the Pre- to During-Olympics, but only of short duration (~47 days), were associated with improvements in biomarkers of several hypothesized pathophysiologic mechanisms (i.e. systemic inflammation, coagulation, pulmonary function, vascular function, autonomic function, oxidative stress, etc.) of air pollution cardio-respiratory health effects. Each subject provided measurements of each biomarker twice before, twice during, and twice after the Olympics. They reported generally large reductions/improvements in most biomarkers (e.g. −60% exhaled nitric oxide [eNO; pulmonary inflammation], −58% urinary 8-OHdG [oxidative stress]; −28% urinary malonaldehyde [oxidative stress]; −34% p-selectin [coagulation], −30% exhaled breath condensate nitrite [pulmonary inflammation]), but not those of heart rate variability.[44,50-52]

Concurrently and using a similar panel study approach, Mu et al (2014) reported increases in peak expiratory flow (measure of lung function) in both men (56 L/min; 13.4%; 95% CI = 9.5%, 14.3%) and women (50 L/min; 17.2%; 95% CI = 12.1%, 22.4%), but not blood pressure in healthy subjects living in Beijing.[53] In a panel study of children living in Beijing, Lin et al (2011) reported reductions in both air pollutant concentrations and exhaled nitric oxide levels during the Olympic Games compared to the Pre-Olympic period, with each 4.0 μg/m^3 increase in black carbon associated and 149 μg/m^3 increase in $PM_{2.5}$ concentration associated with 16.6% (95% CI = 14.1%, 19.2%) and 18.7% (95% CI = 15.0, 22.5%) reductions in eNO, respectively.[54] Further, they reported similar reductions in markers of oxidative stress during the Olympic Games (8-oxodG: −37%, 95% CI = −54%, −16%; urinary malonaldehyde: −25%, 95% CI = −34%, −15%) and similarly, decreased 9-oxodG and malonaldehyde levels associated with increased pollutant concentrations (black carbon, $PM_{2.5}$, SO_2, NO_2, and CO).[55]

Strengths and Limitations—As described by Dominici and Mittleman (2012) in a commentary on Rich et al (2012)[46], these types of 'natural experiment' or accountability studies are particularly compelling for several reasons, including the prospective design with multiple health measurements on each subject during each of the three study periods (Pre-, During-, and Post-Olympic), control of long term time trends in air pollution and health by virtue of this design, the assessment of biomarkers of physiologic mechanisms rather than just cause-specific morbidity and mortality rates, and the ability to assess health effects associated with a reduction in the entire air pollution mixture and not just one pollutant/component.[56] However, this may also be a limitation if you are interested in assessing health responses to one pollutant (e.g. nitrogen dioxide) or pollutant source (e.g. traffic pollution). An important potential limitation of this study, and perhaps of other studies where a regional/city-wide air pollution reduction was associated with a large scale event visible to the population such as the Olympics, is a concomitant change in personal activities. However, based on subject diaries for the 24 hours before each of the 6 clinic visits, they found little change in activity other than small changes in walking/biking and time spent outdoors per day. Thus, changes in personal activities did not appear to explain study finding.[44,46,51]

An additional strength of the Beijing Olympics natural experiment was that other studies evaluated cardiorespiratory outcomes and mortality. For example, Su et al (2015) reported both substantial reductions in pollutant concentrations during the Olympics compared to before the Olympics, and using Quasi-Poisson regression models for the entire study period (May 20 to December 1, 2008), a 9% increase (95% CI = 3%, 15%) in cardiovascular disease mortality (in Beijing residents) associated with each interquartile range increase in ultrafine particle count/concentration (particles <0.1 μm) in the previous 5 days.[57] He et al (2016) similarly reported a 8.4% decrease in all-cause mortality associated with each 10 $\mu g/m^3$ decrease in PM_{10} concentrations during the 2008 Beijing Summer Olympics, and concluded that results were primarily driven by cardiovascular, cerebrovascular, and respiratory mortality.[58] He et al (2016) also reported reduced monthly mortality counts during the Olympics compared to the years before and after the games.[58] Li et al (2010) reported a 46% (95% CI = (−61%, −25%) reduced risk of outpatient asthma visits during the Olympic games (August 8 – September 20, 2008) compared to the a pre-Olympic period (July 1 – August 7, 2008).[59] Taken together, these studies provide strong evidence for effects of air pollution on cardiorespiratory health.

2008 Beijing Summer Olympics – Birth outcomes

Assessments of whether the Beijing Olympics air pollution reductions impacted non-cardiorespiratory outcomes were also conducted. Rich et al (2015) used a birth registry dataset including data from mothers living in 4 central Beijing districts and the same ambient $PM_{2.5}$, NO_2, SO_2, and CO concentrations of Zhang et al (2013) to examine whether the air pollution improvements during the Olympics provided an improvement in fetal growth (using birth weight in term births [≥37 weeks gestational age] as a proxy) among Beijing residents. They reported that babies born to pregnant women living in Beijing whose 8^{th} month of pregnancy was during the 47 day period of lower pollution levels during the Olympics, were 23 g (95% CI = 5g, 40g) larger than babies born to pregnant women with their 8^{th} month of pregnancy during the same calendar dates in 2007 or 2009 (when air pollution levels were higher). Further, increases in multiple pollutants during the 8^{th} month of pregnancy were associated with significant 17g to 34g birth weight reductions in these same children. They did not observe any consistent associations for gestational months 2 to 7.[60]

In other analyses, however, they observed no significant reduction in hypertensive complications of pregnancy (gestational hypertension, preeclampsia, eclampsia) or fetal/placental complications (fetal macrosomia, fetal distress, placental abruption, threatened preterm labor, polyhydramnios, oligohydramnios, or premature rupture of membranes) when any portion of a pregnancy was during the 2008 Olympic Games compared to the same calendar dates in 2009.[61] Huang et al (2015) conducted a similar study in Beijing during the Olympics, using similar birth registry data and ambient air pollution measurements made at multiple PM_{10} and NO_2 monitoring stations across Beijing. Consistent with Rich et al (2015), they reported that among full term births (≥37 weeks gestational age) each 10 μg/m^3 increase in ambient NO_2 concentrations during the third trimester was associated with a 14g reduction in birth weight (95% CI = −21g, −6g) during the third trimester.[62]

Strengths and Limitations—Although a strength of these studies was the use of existing large birth registries providing ample statistical power, they are limited in that each mother's personal activity levels, physical location in and out of Beijing during their pregnancy, and data on biomarker measures of pathophysiologic mechanisms thought to underlie effects of air pollution on fetal growth throughout pregnancy are not known. Further, there was no simultaneous temporal comparison of birth weight in a similar Chinese community not impacted by the Beijing Olympics' air pollution reduction. These studies and that of Parker et al (2008)[16] are the first such natural experiment studies examining whether reduced community levels of air pollutants resulted in improvements in pregnancy outcomes. Together with the first set of Beijing Olympic studies examining biomarker responses described above, they highlight the potential to use a future natural experiment to study potential mechanisms of air pollution medicated fetal growth restriction and other adverse pregnancy outcomes.

Seoul Asian Games

Lee et al (2007) examined whether the Seoul Asian Games in 2002 and its 14 days of traffic volume control in Busan, South Korea reduced air pollutant concentrations and concurrently reduced childhood asthma hospitalizations.[63] Compared to the Pre- and During-Games period, they reported that during the Post-games period, there were 1%-25% reductions in pollutant concentrations (e.g. Pre- and During-Games O_3: 30.3 ppb and PM_{10}: 63.8 μg/m^3; Post-Games O_3: 40.4 ppb and PM_{10}: 64.3μg/m_3), and a reduced risk of childhood asthma hospitalization (RR = 0.73; 95% CI = 0.49, 1.11). Moreover, there was no such pattern in air pollution changes across these same calendar dates in the 3 previous years or 1 subsequent year. Further, using time-series methods, interquartile range increases in PM_{10}, SO_2, NO_2, and O_3 concentrations were associated with significant 24%-35% increases in the risk of hospitalization for asthma, respectively.[64]

Strengths and Limitations—Strengths of the study include the comparison of change patterns in 2002 (year of Asian games) to the change patterns in the 3 years before and 1 year after, thereby controlling for long term time trends. Further this study assessed both changes in an outcome after the games compared to before and during the games, and separately, changes in an outcome associated with individual pollutants,[64] providing more convincing evidence that air pollution reductions were driving the benefit in childhood asthma hospitalizations. However, the duration of the Asian games was only 14 days. Thus an assessment of changes in asthma hospitalizations from the Pre-games to the During-game period only was not possible. Further, there was no control population to determine if any health outcome changes were observed in other communities whose air pollution levels were not impacted by the Asian Games.

CONCLUSIONS AND LESSONS LEARNED

These accountability studies provide some evidence of health impacts of ambient air pollution, and perhaps the health benefits of air quality actions taken to reduce air pollution over the short, medium, or long term. They have examined health responses to air quality actions directed at traffic pollution,[33,37,42,43] industrial emissions,[11-15,17,29] home heating fuel,[18,22] and both long term reductions in the entire air pollution mixture,[30] and short term temporary reductions ranging from 14 to 47 days.[44,50-52,54,55,59,60,62,64] They have assessed concentration changes and health responses to measured/estimated concentrations of TSP,[28,53] PM_{10},[11-15,38,53,64] $PM_{2.5}$,[30,44,46,51,53,54,59] black carbon,[54] black smoke,[18,21,22] sulfate,[33,44,46,51] other PM size fractions,[44,46,51,53,57] PM components and ions,[44,46,51] NO_2 or NO_x,[30,44,46,51,59,60,64] SO_2,[18,21,22,28,29,33,44,46,51,60,64] CO,[44,46,51,59,60] and O_3.[42-44,46,51,59] Further, some of these air pollutant concentration changes have been large (e.g. Beijing, Dublin), while others were smaller in magnitude (e.g. Atlanta). Further, some have examined effects on mortality,[11,13-15,17,18,22,29,33,57,58] hospital/emergency room admissions or outpatient visits,[41-43,59,64] birth outcomes,[16,29,39,60] and various cardiovascular and respiratory biomarkers.[44,46,51,53,55] Some studies examined potentially susceptible populations of children and newborns.[12,14,16,28-30,32,42,54,55,60-62,64]

However, these studies all had various limitations and strengths, some of which are common across the studies. For example, correctly modelling effects of temperature, relative humidity, and other weather covariates, as well as other temporal and subject specific covariates is difficult but important. If and when another natural experiment occurs allowing an assessment of whether air pollution and health are concomitantly improved in a population, we should design our studies to minimize these limitations and maximize the strengths. Several lessons learned from these studies, and recommendations to address them are the following.

First, confounding by long term time trends in air pollution and health rates is important to address, and if not done properly, will limit our ability to make strong or in some cases even valid conclusions about the health benefits of an air quality action. As a means to better control this confounding, future studies assessing long term changes in health outcomes following an air quality action should, if possible, include health data from a population that experienced the same overall health experiences as the population under study before the air quality action, but did not experience the improvement in air quality at the same time. These studies could also present a comparison of demographic characteristics, cause-specific mortality and morbidity rates, and air pollution levels between a potential control population and the population under study, such as in Beatty and Shimshack (2011)[41] and described in Stuart (2010),[65] that would allow the reader to evaluate the suitability of that control population. Ideally, these characteristics, rates, and air pollution levels should be similar between the 2 populations in the pre-intervention period. As shown by the Dublin, Ireland analyses,[18,22] however, identification of an ideal control population is difficult, and thus, some discussion of the suitability of a control population should be included in any description of a study's findings.

Second, most accountability studies were done retrospectively using already collected health data (e.g. hospital admissions with ICD-9 codes for the diagnoses on which the hospitalization was based). This is clearly not a weakness of these studies; it is in fact cost- and time-efficient, and may merely reflect that the purpose of the study was to assess measurable morbidity/mortality impacts in a specified population. However, such studies could incorporate assessments of air pollution impacts on relevant biomarkers to provide a more complete understanding of the any population health response under study. For example, although numerous studies have examined associations between air pollution and myocardial infarction,[66] recent studies have reported an increased risk of ST-elevation myocardial infarction (MI), but not non-ST-elevation MI, associated with short term increases in ambient air pollution.[67,68] However, characterization of MI type and symptom onset time, needed for such a case-crossover study, are not regularly included in the hospitalization record resulting in a limited health effects analyses. Moreover, as demonstrated by the Beijing Olympics studies, prospective data collection of relevant biomarkers and pollutants in a well-defined study population, when coupled with data on these clinical events and proper control of confounding in both analyses, can allow a more complete understanding of if and how any air pollution reduction impacted the health of a community. For example, if the rate of MI is reduced during an air quality intervention, and levels of relevant biomarkers of pathways thought to underlie such an association (e.g. platelet activation, systemic inflammation) are also improved during the same intervention, then this provides evidence of not only if air pollution had a health effect in the population, but also by what mechanism. For studies of reproductive and neurologic outcomes, such a prospective study design with concurrent assessments of clinical outcomes and mechanistic biomarkers could provide strong evidence of if and how air pollution does or does not affect that outcome. Further, simultaneous assessments of clinical outcomes across a range of disease systems (i.e. cardiovascular, respiratory, neurologic, reproductive, etc.), would provide a more complete picture of health impacts of the air quality intervention in the community.

Third, spatial misalignment,[69,70] may be present in several studies. As discussed by Peng and Bell (2010), health data used in several of the studies described above were collected on a regional basis (i.e. across a city, province, or country). However, air pollution data for these studies were often measured at 1 or more fixed monitoring stations in that city, region, or country, often placed in specific locations due to a specific pollutant source or regularly high pollutant concentrations at that location. Data from this monitoring station(s) were then used as proxies for individual subject’s pollutant exposures. This likely resulted in substantial exposure error, as individual subject’s actual exposures differ appreciably across the area, in part due to the chemical constituents of the air pollution mixture. This may be particularly concerning for studies discussed above that used national/regional health data sets and fixed site pollution monitors.[13,15,18,22,29,33,57,58,60,64] Thus, prospective collection of health and exposure data on the same spatial scale should be used, when possible.

Fourth, most studies used regularly collected air pollution measurements at 1 or more monitors in the study region (e.g. $PM_{2.5}$, PM_{10}, SO_2, and O_3). Those that had additional pollutants measured [28,44,50,51] such as concentrations of ultrafine particles, accumulation mode particles, sulfate, nitrate, elemental carbon, and organic carbon, could provide more understanding of what pollutant sources are impacted the most by the air quality action. However, assessments of health risks associated with PM components[71] may require more sophisticated health analyses. For example, as discussed by Mostofsky et al (2012), health effects associated with individual PM constituents that are more highly correlated with PM mass may appear to be more toxic than those constituents not as well correlated with PM mass, even if in fact they are not truly more toxic. Several modelling methods are presented to allow assessment of health effects associated with specific PM constituents independent of PM mass, with interpretations of coefficients also described.[72] In addition, Rich et al (2013) provided another method to assess the impact of PM components on health, by assessing whether the risk of myocardial infarction associated with increased $PM_{2.5}$ mass concentration was modified by the mass fraction of several PM components.[73] Several ongoing studies, currently funded by the Health Effects Institute as part of their Accountability Studies Program, will, when completed, provide additional, sophisticated methods to assess the changes in health outcomes in a population following 1 or more air quality regulations or actions.

Fifth, natural experiments often are less prone to confounding by factors in traditional cohort or case-control studies (e.g. differences in subject characteristics that are correlated with their exposure and independently associated with the health outcome of interest [age, socioeconomic status operationalized as education level or income level, etc.], or weather factors [temperature, relative humidity, etc.]). However, they may be confounded by whether the intervention that resulted in the air pollution decrease (i.e. Olympic Games, legislative mandate requiring fuel oil composition change) was noticeable by city residents which changed their activity frequency and/or level. For example, investigators of the Atlanta or Beijing Olympic studies had to consider confounding by changes in personal activity level, dietary alterations, stress, income level changes, and population influx/efflux. However, policies that result in air pollution improvements that are not evident to the population may not be prone to this type of confounding. Consideration should be taken to identify these factors, and to design ways to collect data to determine if they are in fact a problem for the study and making inference based on it (e.g. subject diaries).

Sixth, the utility of assessing whether the quantitative results and qualitative inference made in the study are robust to statistical analyses and assumptions in them is paramount. As evidenced by the Dockery re-analysis of the Dublin coal sale ban, only when multiple analyses and several control populations are assessed does one conclude that the respiratory mortality changes are not confounded by time trends. However, due to the same set of sensitivity analyses, conclusions of reductions in cardiovascular and total mortality after the Dublin ban compared to before the ban cannot be confidently made. Further, new and previously developed causal inference methods have been proposed for such accountability studies. [23,74]

Last, it is important to note the difference between natural experiment studies assessing short term (days to weeks; e.g. [11,13,15,42-44,46,51,53,55,57,59,64]) and long term (years; e.g. [18,22,30,33]) relationships between air pollution reductions and any beneficial health effects. Correctly adjusting for background trends in morbidity/mortality and air pollution levels are likely more of an issue in the long term studies. Whereas the use of an A-B-A design to evaluate whether biomarker levels decrease when air pollution decreases (from time A to B), and then increase when air pollution levels increase again (from B to A), is more important in short term studies so as to establish that the biomarker changes are driven by the air pollution changes.

Overall, these accountability studies have provided important assessments of whether environmental policies and individual air quality actions around the world have benefited public health. In some, they have also provided important evidence to assess if and by what mechanisms air pollution may cause a health event or contribute to underlying disease. Future natural experiments may occur, and if so, epidemiologic lessons learned from these studies may inform us as to how to maximize the information learned from the future events. If so, we may gain information about new pollutant/outcome associations, specific diseases impacted or not impacted by air pollution, sources of pollution most strongly linked or not linked at all to health effects, and/or new groups of susceptible individuals requiring protection.

Acknowledgements

This work was funded by grants from the Electric Power and Research Institute, National Institutes of Environmental Health Sciences (R01-ES019165-01), and the New York State Energy Research and Development Authority (Contract #’s 59800, 100412). I would like to thank Drs. Annette Rohr, Ronald Wyzga, Annemoon van Erp, Johanna Boogaard, and Thomas Bateson for review and/or comment on the manuscript.

REFERENCES

1. Brook RD, Franklin B, Cascio W, et al. Air pollution and cardiovascular disease: a statement for healthcare professionals from the Expert Panel on Population and Prevention Science of the American Heart Association. Circulation. 2004; 109:2655–71. [PubMed: 15173049]
2. Brook RD, Rajagopalan S, Pope CA 3rd, et al. Particulate matter air pollution and cardiovascular disease: An update to the scientific statement from the American Heart Association. Circulation. 2010; 121:2331–78. [PubMed: 20458016]
3. United States Environmental Protection Agency (US EPA). NCEA - Office of Research and Development. Integrated Science Assessment for Particulate Matter. 2009.
4. Bell ML, Ebisu K, Leaderer BP, et al. Associations of PM(2).(5) constituents and sources with hospital admissions: analysis of four counties in Connecticut and Massachusetts (USA) for persons >/= 65 years of age. Environ Health Perspect. 2014; 122:138–44. [PubMed: 24213019]
5. Dominici F, Peng RD, Bell ML, et al. Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases. JAMA. 2006; 295:1127–34. [PubMed: 16522832]
6. Assessing health impact of air quality regulations: concepts and methods for accountability research. Health Effects Institute; 2003. HEI Communication 11 [Accessed December 23, 2015]
7. van Erp AM, Kelly FJ, Demerjian KL, Pope CA III, Cohen AJ. Progress in research to assess the effectiveness of air quality interventions towards improving public health. Air Quality Atmosphere and Health. 2012; 5:217–30.
8. Currie J, Ray SH, Neidell M. Quasi-experimental studies suggest that lowering air pollution levels benefits infants' and children's health. Health Aff (Millwood). 2011; 30:2391–9. [PubMed: 22147868]
9. Granslo JT, Bratveit M, Hollund BE, et al. Airway symptoms and lung function in the local population after the oil tank explosion in Gulen, Norway. BMC Pulm Med. 2012; 12:76. [PubMed: 23234609]
10. Thomas PA, Brackbill R, Thalji L, et al. Respiratory and other health effects reported in children exposed to the World Trade Center disaster of 11 September 2001. Environ Health Perspect. 2008; 116:1383–90. [PubMed: 18941582]

11. Pope CA 3rd. Respiratory disease associated with community air pollution and a steel mill, Utah Valley. Am J Public Health. 1989; 79:623–8. [PubMed: 2495741]
12. Ransom MR, Pope CA 3rd. Elementary school absences and PM10 pollution in Utah Valley. Environ Res. 1992; 58:204–19. [PubMed: 1511674]
13. Pope CA 3rd. Respiratory hospital admissions associated with PM10 pollution in Utah, Salt Lake, and Cache Valleys. Arch Environ Health. 1991; 46:90–7. [PubMed: 2006899]
14. Pope CA III, Dockery DW. Acute health effects of PM10 pollution on symptomatic and asymptomatic children. American Review of Respiratory Disease. 1992; 145:1123–6. [PubMed: 1586057]
15. Pope CA 3rd, Schwartz J, Ransom MR. Daily mortality and PM10 pollution in Utah Valley. Arch Environ Health. 1992; 47:211–7. [PubMed: 1596104]
16. Parker JD, Mendola P, Woodruff TJ. Preterm birth after the Utah Valley Steel Mill closure: a natural experiment. Epidemiology. 2008; 19:820–3. [PubMed: 18854706]
17. Pope CA 3rd, Rodermund DL, Gee MM. Mortality effects of a copper smelter strike and reduced ambient sulfate particulate matter air pollution. Environ Health Perspect. 2007; 115:679–83. [PubMed: 17520052]
18. Clancy L, Goodman P, Sinclair H, Dockery DW. Effect of air-pollution control on death rates in Dublin, Ireland: an intervention study. Lancet. 2002; 360:1210–4. [PubMed: 12401247]
19. Flanagan, PJ. Air quality in Ireland: the present position. An Foras Forbartha; Dublin: 1986.
20. Kelly I, Clancy L. Mortality in a general hospital and urban air pollution. Irish Journal of the Medical Society. 1984; 77:322–4.
21. Goodman PG, Rich DQ, Zeka A, Clancy L, Dockery DW. Effect of sequential air pollution controls on black smoke and sulphur dioxide concentrations across Ireland. Journal of the Air and Waste Management Association. 2009; 59:207–13.
22. Dockery DW, Rich DQ, Goodman PG, et al. Effect of air pollution control on mortality and hospital admissions in Ireland. Res Rep Health Eff Inst. 2013:3–109.
23. Cox LA Jr. Popken DA. Has reducing fine particulate matter and ozone caused reduced mortality rates in the United States? Annals of Epidemiology. 2015; 25:162–73. [PubMed: 25571792]
24. Umweltbundesamt. Daten zur Umwelt 1995/96. Erich Schmidt Verlag; Berlin: 1997.
25. Pitz M, Kreyling WG, Hoelscher B, Cyrys J, Wichmann HE, Heinrich J. Change of the ambient particle size distribution in East Germany between 1993 and 1999. Atmospheric Environment. 2001; 109:325–33.
26. Ebelt S, Brauer M, Cyrys J, et al. Air quality in postunification Erfurt, East Germany: associating changes in pollutant concentrations with changes in emissions. Environ Health Perspect. 2001; 109:325–33.
27. Wichmann HE, Spix C, Tuch T, et al. Daily mortality and fine and ultrafine particles in Erfurt, Germany part I: role of particle number and particle mass. Res Rep Health Eff Inst. 2000:5–86. discussion 7-94.
28. Heinrich J, Hoelscher B, Frye C, et al. Improved air quality in reunified Germany and decreases in respiratory symptoms. Epidemiology. 2002; 13:394–401. [PubMed: 12094093]
29. Luechinger S. Air pollution and infanta mortality: a natural experiment from power plant desularization. Journal of Health Economics. 2014; 37:219–31. [PubMed: 25105867]
30. Gauderman WJ, Urman R, Avol E, et al. Association of improved air quality with lung development in children. N Engl J Med. 2015; 372:905–13. [PubMed: 25738666]
31. Dockery DW, Ware JH. Cleaner air, bigger lungs. N Engl J Med. 2015; 372:970–2. [PubMed: 25738674]
32. Peters J, Hedley AJ, Wong CM, et al. Effects of an ambient air pollution intervention and environmental tobacco smoke on children's respiratory health in Hong Kong. Int J Epidemiol. 1996; 25:821–8. [PubMed: 8921462]
33. Hedley AJ, Wong CM, Thach TQ, Ma S, Lam TH, Anderson HR. Cardiorespiratory and all-cause mortality after restrictions on sulphur content of fuel in Hong Kong: an intervention study. Lancet. 2002; 360:1646–52. [PubMed: 12457788]

34. Wong CM, Lam TH, Peters J, et al. Comparison between two districts of the effects of an air pollution intervention on bronchial responsiveness in primary school children in Hong Kong. J Epidemiol Community Health. 1998; 52:571–8. [PubMed: 10320858]

35. London, Tf. Fourth annual report. Transport for London; London: 2006. Central London.

36. Kelly F, Armstrong B, Atkinson R, et al. The London low emission zone baseline study. Res Rep Health Eff Inst. 2011:3–79.

37. Tonne C, Beevers S, Armstrong B, Kelly F, Wilkinson P. Air pollution and mortality benefits of the London Congestion Charge: spatial and socioeconomic inequalities. Occup Environ Med. 2008; 65:620–7. [PubMed: 18308748]

38. Kelly F, Anderson HR, Armstrong B, et al. The impact of the congestion charging scheme on air quality in London. Part 1. Emissions modeling and analysis of air pollution measurements. Res Rep Health Eff Inst. 2011:5–71.

39. Currie J, Walker R. Traffic Congestion and Infant Health: Evidence from E-ZPass. American Economic Journal: Applied Economics. 2011; 3:65–90.

40. Adar SD, D'Souza J, Sheppard L, et al. Adopting Clean Fuels and Technologies on School Buses. Pollution and Health Impacts in Children. Am J Respir Crit Care Med. 2015; 191:1413–21. [PubMed: 25867003]

41. Beatty TK, Shimshack JP. School buses, diesel emissions, and respiratory health. J Health Econ. 2011; 30:987–99. [PubMed: 21741102]

42. Friedman MS, Powell KE, Hutwagner L, Graham LM, Teague WG. 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. [PubMed: 11180733]

43. Peel JL, Klein M, Flanders WD, Mulholland JA, Tolbert PE, Committee HEIHR. Impact of improved air quality during the 1996 Summer Olympic Games in Atlanta on multiple cardiovascular and respiratory outcomes. Res Rep Health Eff Inst. 2010:3–23. discussion 5-33.

44. Zhang J, Zhu T, Kipen H, et al. Cardiorespiratory biomarker responses in healthy young adults to drastic air quality changes surrounding the 2008 Beijing Olympics. Res Rep Health Eff Inst. 2013:5–174.

45. Zhu T, Wang M, Zheng J, et al. Use of a mobile laboratory to evaluate changes in on-road air pollutants during the Beijing 2008 Summer Olympics. Atmos Chem Phys. 2009; 9:8247–63.

46. Rich DQ, Kipen HM, Huang W, et al. Association between changes in air pollution levels during the Beijing Olympics and biomarkers of inflammation and thrombosis in healthy young adults. JAMA. 2012; 307:2068–78. [PubMed: 22665106]

47. Wang M, Zhu T, Zheng J, et al. Use of a mobile laboratory to evaluate changes in on-road air pollutants during the Beijing 2008 Summer Olympics. Atmos Chem Phys. 2009; 9:8247–63.

48. Wang T, Nie W, Gao J, et al. Air quality during the 2008 Beijing Olympics: secondary pollutants and regional impact. Atmos Chem Phys. 2010; 10:7603–15.

49. Wang T, Xie SD. Assessment of traffi c-related air pollution in the urban streets before and during the 2008 Beijing Olympic Games traffi c control period. Atmospheric Environment. 2009; 43:5682–90.

50. Rich DQ, Kipen HM, Zhang J, Kamat L, Wilson AC, Kostis JB. Triggering of transmural infarctions, but not nontransmural infarctions, by ambient fine particles. Environ Health Perspect. 2010; 118:1229–34. [PubMed: 20435544]

51. Huang W, Wang G, Lu SE, et al. Inflammatory and oxidative stress responses of healthy young adults to changes in air quality during the Beijing Olympics. Am J Respir Crit Care Med. 2012; 186:1150–9. [PubMed: 22936356]

52. Gong J, Zhu T, Kipen H, et al. Malondialdehyde in exhaled breath condensate and urine as a biomarker of air pollution induced oxidative stress. J Expo Sci Environ Epidemiol. 2013; 23:322–7. [PubMed: 23321859]

53. Mu L, Deng F, Tian L, et al. Peak expiratory flow, breath rate and blood pressure in adults with changes in particulate matter air pollution during the Beijing Olympics: a panel study. Environ Res. 2014; 133:4–11. [PubMed: 24906062]

54. Lin W, Huang W, Zhu T, et al. Acute respiratory inflammation in children and black carbon in ambient air before and during the 2008 Beijing Olympics. Environ Health Perspect. 2011; 119:1507–12. [PubMed: 21642045]
55. Lin W, Zhu T, Xue T, et al. Association between changes in exposure to air pollution and biomarkers of oxidative stress in children before and during the Beijing Olympics. Am J Epidemiol. 2015; 181:575–83. [PubMed: 25770981]
56. Dominici F, Mittleman MA. China's air quality dilemma: reconciling economic growth with environmental protection. JAMA. 2012; 307:2100–2. [PubMed: 22665112]
57. Su C, Hampel R, Franck U, et al. Assessing responses of cardiovascular mortality to particulate matter air pollution for pre-, during- and post-2008 Olympics periods. Environ Res. 2015; 142:112–22. [PubMed: 26133808]
58. He G, Fan M, Zhou M. The effect of air pollution on mortality in China: evidence from the 2008 Beijing Olympic Games. Journal of Environmental Economics and Management. 2016; 2016:18–39.
59. Li Y, Wang W, Kan H, Xu X, Chen B. Air quality and outpatient visits for asthma in adults during the 2008 Summer Olympic Games in Beijing. Sci Total Environ. 2010; 408:1226–7. [PubMed: 19959207]
60. Rich DQ, Liu K, Zhang J, et al. Differences in Birth Weight Associated with the 2008 Beijing Olympic Air Pollution Reduction: Results from a Natural Experiment. Environ Health Perspect. 2015
61. Assibey-Mensah V, Liu K, Thurston SW, et al. Impact of the 2008 Beijing Olympics on the Risk of Pregnancy Complications. Arch Environ Occup Health. 2015 0.
62. Huang C, Nichols C, Liu Y, et al. Ambient air pollution and adverse birth outcomes: a natural experiment study. Population Health Metrics. 2015; 13:17. [PubMed: 26190943]
63. Lee JT, Son JY, Cho YS. Benefits of mitigated ambient air quality due to transportation control on childhood asthma hospitalization during the 2002 summer Asian games in Busan, Korea. J Air Waste Manag Assoc. 2007; 57:968–73. [PubMed: 17824287]
64. Lee JT, Son JY, Cho YS. Benefits of miitigated ambient air quality due to transportation control on childhood asthma hospitalization during the 2002 Summer Asian Games in Busan, Korea. Journal of the Air & Waste Management Association. 2012; 57:968–73.
65. Stuart EA. Matching methods for causal inference: A review and a look forward. A Review Journal of the Institute of Mathematical Statistics. 2010; 25:1–21.
66. Mustafic H, Jabre P, Caussin C, et al. Main air pollutants and myocardial infarction: a systematic review and meta-analysis. JAMA. 2012; 307:713–21. [PubMed: 22337682]
67. Gardner B, Ling F, Hopke PK, et al. Ambient fine particulate air pollution triggers ST-elevation myocardial infarction, but not non-ST elevation myocardial infarction: a case-crossover study. Part Fibre Toxicol. 2014; 11:1. [PubMed: 24382024]
68. Pope CA, Muhlestein JB, Anderson JL, et al. Short-Term Exposure to Fine Particulate Matter Air Pollution Is Preferentially Associated With the Risk of ST-Segment Elevation Acute Coronary Events. J Am Heart Assoc. 2015; 4
69. Peng RD, Bell ML. Spatial misalignment in time series studies of air pollution and health data. Biostatistics. 2010; 11:720–40. [PubMed: 20392805]
70. Gryparis A, Paciorek CJ, Zeka A, Schwartz J, Coull BA. Measurement error caused by spatial misalignment in environmental epidemiology. Biostatistics. 2009; 10:258–74. [PubMed: 18927119]
71. Ito KMR, Ross Z, Nadas A, Thurston G, Matte T. Fine particulate matter constituents associated with cardiovascular hospitalizations and mortality in New York City. Environ Health Perspect. 2001:467–73.
72. Mostofsky E, Schwartz J, Coull BA, et al. Modeling the association between particle constituents of air pollution and health outcomes. Am J Epidemiol. 2012; 176:317–26. [PubMed: 22850792]
73. Rich DQ, Ozkaynak H, Crooks J, et al. The triggering of myocardial infarction by fine particles is enhanced when particles are enriched in secondary species. Environ Sci Technol. 2013; 47:9414–23. [PubMed: 23819750]

74. Zigler CM, Kim C, Choirat C, et al. Causal Inference Methods for Estimating Long-Term Health Effects of Air Quality Regulations. Res Rep Health Eff Inst. 2016:5–49. [PubMed: 27526497]

75. Samet, JM., Cohen, AJ., Demerjian, KL., et al. Assessing Health Impact of Air Quality Regulations: Concepts and Methods for Accountability Researc. Health Effects Institute; Boston, MA: 2003. Communication 11

Highlights

1. Natural experiments assess if regulatory actions resulted in lower air pollution
2. Natural experiments also assess if beneficial health effects were observed
3. Strengths and limitations of past designs can inform future study designs

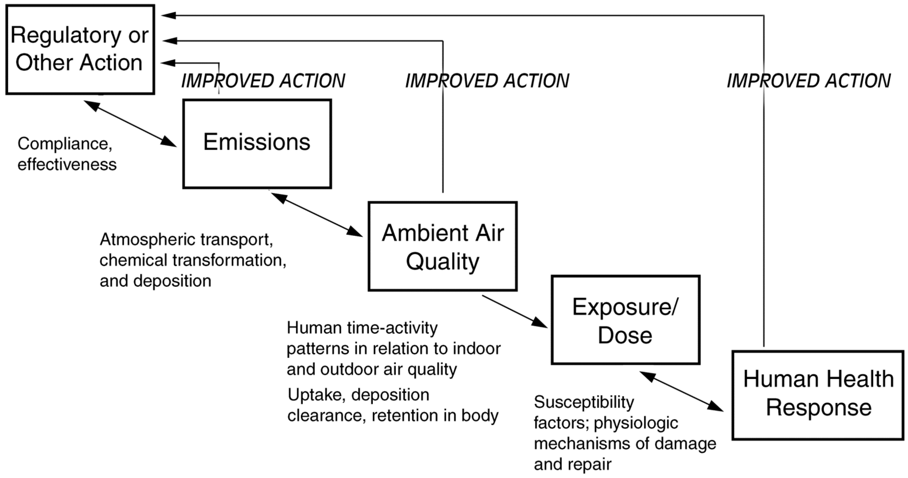

Figure 1.

Chain of Accountability. *Source: Health Effects Institute*. [75]

Table 1

Characteristics, research questions, and study findings, strengths and limitations of each natural experiment.

Study and Location	Description of Natural Experiment	Chain of Accountability Question	Study Objective(s)	Main Findings	Major Strengths	Major Limitations
Pope et al. 1989, 1991, 1992a, 1992b; Ransom et al. 1992 Utah Valley, Utah	Closure of a large steel mill from August 1986 to September 1987	Did reduced emissions lead to improved air quality and a beneficial human health response?	• To compare PM_{10} concentrations and the number of health events during strike to before and after strike	Average PM_{10} concentration reduced from 90 μg/m^3 to 51 μg/m^3 During strike, reduced: • hospital admissions for pneumonia, pleurisy, bronchitis, and asthma, • childhood school absenteeism, • bronchitis and asthma admissions for pre-school aged children, • respiratory symptoms and reduced peak expiratory flow, • total non-accidental mortality, respiratory mortality, and cardiovascular mortality	• Use of already collected air pollution and hospital admissions and mortality data • A-B-A design to control for confounding by time trends in mortality/morbidity	• Did not have simultaneous temporal comparisons in other Utah counties (control counties) to provide additional control for confounding by time trends
Parker et al, 2008 Utah Valley and other counties in Utah	Closure of a large steel mill from August 1986 to September 1987;	Did reduced emissions lead to a beneficial human health response?	To compare the preterm birth rate during the strike to before and after strike, both in counties in the Utah Valley, and other Utah counties	• Utah Valley mothers who were already pregnant at the time of mill closure were less likely to deliver prematurely than mothers pregnant before or after the closure. • No pattern in mothers from other Utah counties	• Use of already collected birth data • A-B-A design to control for confounding by time trends • Simultaneous temporal comparison in other Utah counties (control counties) provides additional control for time trends	• 13 month period too long to examine specific windows of pregnancy • No examination of air pollution changes in other counties • No complimentary analysis of air pollution and preterm birth to confirm that air pollution changes drove changes in Utah Valley preterm birth rate
Pope et al, 2007 New Mexico, Arizona, Utah, and Nevada	Nationwide copper smelter strike from July 1967 to April 1968	Did reduced emissions lead to a beneficial human health response?	• Compared monthly mortality rates during 8.5 months of smelter strike to rates before	• Estimated reduction of suspended sulfate of ~2.5 μg/m^3 • 1.5% - 4.0% decrease in mortality during strike compared	• Use of already collected mortality data • A-B-A design to control for confounding by time trends	• No complementary estimation of change in mortality associated with sulfate or other

Study and Location	Description of Natural Experiment	Chain of Accountability Question	Study Objective(s)	Main Findings	Major Strengths	Major Limitations
			and after strike (1960 – 1975)		• Adjustment for regional and national cardiovascular and respiratory mortality rates to control for confounding by background trends in cardiorespiratory mortality • Adjustment for nationwide influenza/pneumonia rates to control for confounding by national or regional epidemics • Adjusted for monthly mortality counts in neighboring states (control states) to account for confounding by time trends.	pollutant concentrations
Clancy et al, 2002 Dublin, Ireland	Ban on marketing and sale of coal in Dublin starting in September, 1990	Did reduced emissions lead to improved air quality and a beneficial human health response?	•Compared black smoke and sulfur dioxide concentrations and cause specific mortality rates in the 6 years before and after the ban	Black smoke reductions: • 50.2 to 14.6 μg/m^3 in all seasons • 85.4 to 21.5 μg/m^3 in winter Sulfur dioxide reductions: • 33.4 to 22.1 μg/m^3 in all seasons • 40.4 to 24.9 μg/m^3 in winter Estimated mortality reductions: • total non-trauma: −6% • respiratory: −16% • cardiovascular: −10%	• Use of nationally-maintained mortality data • Use of already collected black smoke and sulfur dioxide measurements • Adjustment for influenza epidemics to control for other causes of respiratory mortality	• Adjustment for national cause-specific mortality rates, does not completely control for background mortality rate changes, resulting in residual confounding by time trends • Population size and age distribution changed substantially during study period. Could only estimate population size, potentially resulting in bias
Dockery et al, 2013 and Goodman et al, 2009 • Cork, Ireland • Arklow, Drogheda, Dundalk,	Ban on marketing and sale of coal in • Cork (1995) • 5 cities (1998)	Did reduced emissions lead to improved air quality and a beneficial human health	• To compare cause specific mortality rates and hospital admissions in the 5 years	Black smoke reductions (all seasons): • Cork: −49%, 33.7 to 17.2 μg/m^3 • 5 Cities: −48% to −61% • No reductions in	• Use of nationally-maintained mortality data for 10 years rather than 6 years, resulting in better control for time	• Population size and age distribution changed substantially during study

Study and Location	Description of Natural Experiment	Chain of Accountability Question	Study Objective(s)	Main Findings	Major Strengths	Major Limitations
Limerick, & Wexford, Ireland • Dublin, Ireland (reanalysis)	• Dublin (1990, reanalysis)	response?	before and after the ban in Cork, 5 Cities, and 12 Midland Counties (as a comparison) • To compare cause specific mortality rates in the 10 years before and after the ban in Dublin (re-analysis)	total and cardiovascular mortality in any city • Significant reductions in respiratory mortality in Dublin (−17%), and non-significant reductions in Cork (−9%) and 5 Cities (−3%) • Dublin ban associated with non-significant small changes (−2% to −3%) in cause-specific mortality rates in 12 Midland counties	trends • Use of already collected black smoke measurements • Adjustment for influenza epidemics to control for other causes of respiratory mortality • Adjustment for mortality rates in coastal counties rather than entire country likely better control of confounding than Clancy et al (2002) • Assessment of 12 Midland Counties as a comparison	period. Could only estimate population size, potentially resulting in bias
Heinrich et al, 2002; Former East Germany	German reunification	Did improved air quality lead to a beneficial human health response?	• Cross-sectional surveys of children in 3 phases to assess prevalence of respiratory disorders • Estimates of change in respiratory disorders associated with changes in TSP and SO_2	• Each 50 μg/m^3 increase in TSP associated with increased prevalence of: • bronchitis (202%) • sinusitis (158%) • frequent colds (90%) • febrile infections (79%) • Each 100 μg/m^3 increase in SO_2 associated with increased prevalence of: • bronchitis (172%) • sinusitis (126%) • frequent colds (81%) • febrile infections (76%)	• Recruited only children living in study areas for 2+years and who had not relocated from more than 2km away • Prospective health data collection should minimize misclassification • Comparison of same children across study periods • Adjustment for multiple measures of socioeconomic status and other environmental contaminants	• Potential for residual confounding by a change to western lifestyle a • Potential for residual confounding by time trends due to no control population experiencing same health rates, but not air pollution changes
Luechinger et al, 2014 Germany	Mandated scrubber installations at all power plants	Did reduced emissions lead to a beneficial human health response?	• Using national mortality and birth data, they estimate effects of this mandated scrubber installations and air pollution	Each 1 μg/m^3 decrease in annual county average SO_2 concentration associated with 2.6 less infant deaths per 100,000 births •Decreased SO_2 concentrations also associated with increased infant lengths and	• Use of already collected air pollution data • Use of already collected birth and infant mortality data for a large population	• No control population to know if similar patterns of lung function improvement observed in areas without an air pollution concentration decrease

Study and Location	Description of Natural Experiment	Chain of Accountability Question	Study Objective(s)	Main Findings	Major Strengths	Major Limitations
			reductions on neonatal mortality and the number of infants with low birth weight and length	weights		• Assessed only 1 pollutant and thus could not evaluate associations between these same outcomes and concurrent changes in the concentrations of other pollutants
Gauderman et al, 2015 Long Beach, Mira Loma, Riverside, San Dimas, and Upland, California	Air quality control policies targeted at mobile and stationary sources have reduced air pollutant levels over past few decades in California	Did improved air quality lead to a beneficial human health response?	• Assessment of changes in lung function from ages 11 to 15 associated with decreases in air pollutants during the same time, in children from 3 separate cohorts corresponding to 3 separate calendar periods (1994-2011), as part of the Children's Health Study	• Decreased NO_2, $PM_{2.5}$, and PM_{10} concentrations associated with increases in FEV_1 and FVC in children from 11 to 15 years of age	• Large, well characterized cohort with prospective pulmonary function measurements • Use of available air pollution data • Within community analyses, presumably resulting in no confounding by differences between communities	• Potential residual confounding by changes in racial/ethnic composition and other population characteristics within a community over time • No control population possible, to know if similar patterns of lung function improvement observed in areas without an air pollution concentration decrease
Hedley et al, 2002 Hong Kong	Reduction in sulfur content of fuel oil used in power plants and on-road vehicles from 1990-1995	Did a regulatory action lead to improved air quality and a beneficial human health response?	• To compare changes in monthly deaths and monthly mean air pollutant concentrations in Hong Kong between 1985 and 1995	• Reductions from before to after intervention in SO_2 (mean change = −45% over 5 years after intervention) • Seasonal reductions in respiratory, cardiovascular, and total mortality after the intervention	• Use of already collected air pollution and mortality data	• Although likely minimal, residual confounding by infectious disease mortality including influenza is possible • A-B design may result in residual confounding by time trends in mortality and air pollution • No control population could be used to determine if similar changes in seasonal patterns of mortality were observed

Study and Location	Description of Natural Experiment	Chain of Accountability Question	Study Objective(s)	Main Findings	Major Strengths	Major Limitations
						elsewhere
Kelly et al, 2011 London, England	In 2008, low emission zone established in London (2644 km^2) to restrict entry of old and polluting diesel vehicles, but not cars or motorcycles	Did a regulatory action lead to improved air quality?	• To determine whether reductions in congestion, traffic volume, and air pollution concentrations within London zone (2644 km^2) were observed • To assess whether a primary care dataset in London could be used to assess whether any reduction in traffic and air pollution had a health benefit	• 26% reduction in the excess delay per kilometer • Reduced traffic volume during the charging hours compared to before the charging scheme was implemented • Modeling efforts predicted 3.8% reductions in NO_x and −2.6% reductions in PM_{10} emissions after 2 years • Pilot analyses found no association between NOx concentrations and prevalence of respiratory outcomes	• Use of both traffic and air pollution data to assess environmental impacts of policy • Use of primary care dataset to provide comprehensive health data for the population of interest	• Some primary care providers were concerned that patient confidentiality could not be certain and thus did not make some health records available for inclusion in the study, and thus statistical power for future analyses of any health benefits of the low emission zone may be reduced
Currie and Walker, 2011 New Jersey and Pennsylvania, neighborhoods near highway toll plazas	E-ZPass, an electronic toll collection system not requiring cars to stop at toll plazas, was installed at toll plazas in New Jersey and Pennsylvania from 1997 to 2000	Did a policy lead to a beneficial health response?	• To determine whether implementation of E-ZPass resulted in improved birth outcomes among women living near the toll plazas	• Introduction of E-ZPass reduced preterm births (<37 weeks gestation) by 10.8% and low birth weight (<2500g) by 11.8% among mothers living within 2 km of a toll plaza, relative to mothers living 2-10km from a toll plaza	• Use of existing birth records/data from both states • Use of a control population to allow comparison of changes in the frequency of preterm birth and low birth weight before and after E-ZPass, both in mothers living within 2 km of a toll plaza, and mothers living 2-10km from a toll plaza	• No air pollution monitoring done to determine the change in air pollutant concentrations after E-ZPass installation in areas near the toll plazas (<2km) and farther away from the toll plazas (2-10 km) to confirm E-ZPass program is driving health benefit • No available air pollution monitoring data at toll plazas to assess air pollutant concentration changes there
Beatty and Shimshack, 2011 Puget Sound area of Washington	Localized emissions reduction	Do regulatory actions result in a beneficial	• To determine whether child and adult	• School bus retrofits induced large reductions in monthly counts of	• Use of a large existing statewide dataset on hospital	• No air pollution monitoring done to determine the

Study and Location	Description of Natural Experiment	Chain of Accountability Question	Study Objective(s)	Main Findings	Major Strengths	Major Limitations
State	program in Washington state to retrofit diesel school buses with aggressive pollution control technologies	health response?	residents of school districts where diesel bus retrofits were adopted had a greater reduction in monthly counts of hospital admissions for bronchitis, asthma, pneumonia, and pleurisy after the retrofits compared to before the retrofits, relative to residents of the non-adopter school districts	respiratory hospital admissions for both children and adults with chronic conditions • Adopter school districts experience 23% and 37% fewer admissions for bronchitis/asthma and pleurisy/pneumonia in children, relative to non-adopter school districts	admissions likely captured nearly all health outcomes requiring hospitalization in this population • Use of a control population to allow comparison of changes in respiratory outcomes from before to after the retrofits, both in residents of school districts adopting and not-adopting the diesel bus retrofits for their school buses • Several sensitivity analyses conducted to assess whether the control population (non-adopter districts) were similar to adopter districts with regard to demographics and health outcomes before the retrofit program was begun	change in air pollutant concentrations after school bus retrofits in adopter districts relative to non-adopter districts
Friedman et al, 2001 Atlanta, Georgia	1996 Atlanta Olympics – policies to reduce vehicular traffic and congestion	Did a regulatory action lead to improved air quality and a beneficial human health response?	• To determine whether there were reductions in hospital, emergency room, and urgent care center visit data for children (aged 1-5 years) residents of 5 metropolitan Atlanta counties during the games • To determine whether there were reductions in central site	• Reduced peak daily O_3 concentrations (81.3 ppb before games to 58.6 ppb during games), and 22.5% reduction in morning traffic counts during Olympics compared to 4 weeks before and after games • 41.6% reduction in acute care and hospitalizations for asthma during Olympic Games (4.23 to 2.47 daily events)	• Use of existing hospital and emergency room data • Use of existing air pollution data • A-B-A study design to control for confounding by time trends	• Only 73 day study period and only 17 day period of Olympics gave limited statistical power • Potential for residual confounding by seasonal patterns in pollution and hospitalizations • Potential for residual confounding by Atlanta residents changing behaviors and personal activity levels during Games

Study and Location	Description of Natural Experiment	Chain of Accountability Question	Study Objective(s)	Main Findings	Major Strengths	Major Limitations
			PM_{10}, CO, and O_3 measurements from the same 5 counties during the games			
Peel et al, 2010 Atlanta, Georgia	1996 Atlanta Olympics – policies to reduce vehicular traffic and congestion	Did a regulatory action lead to improved air quality and a beneficial human health response?	• To determine whether there were reductions in hospital and ER admissions data from 1995 to 2004 • To determine whether there were reductions in air pollutant concentrations (PM_{10}, 8-hour maximum O_3, 1 hour maximum O_3, NO_2, NO_x, SO_2, and CO) measured at stations both within and outside Atlanta • Used traffic data from 18 sites within the 5 county area	• 20%-30% reductions in ambient ozone concentrations in Atlanta, but similar declines in ozone concentration were observed throughout Georgia and the southeastern United States during this time • 2% to 20% reductions in weekday peak morning traffic counts, consistent with, but smaller in magnitude[43] than those of Friedman et al (2001). • No reduction in cardiovascular or respiratory ED visits during the Olympics, which was inconsistent with the earlier work by Friedman et al (2001)	Compared to Friedman: • Longer time period with health and air pollution data allowed better control of season and long term time trends) • Traffic data from 18 sites within the five counties better represented regional traffic patterns • More spatially diverse air pollution data Overall: • Large, well characterized ER dataset within a large urban area	• Only 73 day study period and only 17 day period of Olympics gave limited statistical power • Air pollution reduction may have not been large enough to elicit a detectable health response • lack of a 'control' area/county/city with a similar population where such a traffic reduction system was not available, which may result in residual confounding by time trends
Zhang et al, 2013; Rich et al, 2012; Huang et al, 2012 Beijing, China	2008 Summer Olympics – Driving restrictions, industry and construction shutdowns, drastic reductions in emissions from numerous sources	Did a regulatory action lead to improved air quality and a beneficial human health response?	• In a panel study of healthy young medical residents, were the levels of cardio-pulmonary biomarkers and (measured twice before, during, and after Olympics) and multiple air pollutants lower during the games compared to before and after	• 18% to 60% reductions in all pollutants, but ozone (20% increase) • Large reductions in most biomarkers, but not heart rate variability markers.	• A-B-A design to control for time trends • Prospective measurement of all health and pollution data • Measurement of biomarkers allowing investigation of impacts of air pollution on physiologic mechanisms • Assessed numerous pollutants including ions and PM components	• All pollutants, but ozone, were reduced simultaneously making assessment of health responses associated with individual pollutants difficult • These are healthy young subjects, and thus not those where such biomarker changes would be indicative of actual clinical events (e.g. myocardial infarction, stroke)

Study and Location	Description of Natural Experiment	Chain of Accountability Question	Study Objective(s)	Main Findings	Major Strengths	Major Limitations
			the games			
Mu et al, 2014 Beijing, China	2008 Summer Olympics – Driving restrictions, industry and construction shutdowns, drastic reductions in emissions from numerous sources	Did a regulatory action lead to a beneficial human health response?	• In a panel study of healthy adult subjects (20-65 years of age) living in Haidan district of Beijing, were there improvements in peak expiratory flow, blood pressure, and respiration rate, measured once before, once during, and once after the Olympics, as well as PM_1, $PM_{2.5}$, PM_7, PM_{10} and TSP measured at 1 centrally located monitor	• All air pollutant concentrations were 54%-60% lower during the games compared to before the games • Peak expiratory flow increased in 80% of study subjects from before to during the Olympic Games. •Percent of subjects with a fast respiration rate (>20/min) decreased during the games compared to before the games, and increased after the games compare to the games • No clear pattern of blood pressure change across periods	• A-B-A design to control for time trends • Prospective measurement of all health and pollution data • Measurement of biomarkers allowing investigation of mechanism	• Potential for residual confounding by a change in personal activities by study subjects during the games, compared to before and after the games.
Lin et al, 2011 Beijing, China	2008 Summer Olympics – Driving restrictions, industry and construction shutdowns, drastic reductions in emissions from numerous sources	Did a regulatory action lead to improved air quality and a beneficial human health response?	• In a panel study of 36 elementary school aged children (with 5 observation periods for each subject over a 2 year study period, were there lower levels of exhaled nitric oxide (made at multiple times during each observation period) , and black carbon (BC), $PM_{2.5}$, NO_x, SO_2, and CO concentrations (measured at a	• BC and $PM_{2.5}$ concentrations substantially lower during games, compared to before the games • Increases in BC (4.0 $\mu g/m^3$) and $PM_{2.5}$ (149 $\mu g/m^3$) associated with 16.6% and 18.7% increases in exhaled nitric oxide, respectively	• Prospective measurement of all health and pollution data • Measurement of biomarkers allowing investigation of mechanism	• Did not have A-B-A design to control for time trends • Potential for residual confounding by a change in personal activities by study subjects during the games, compared to before and after the games. • Did not directly estimate change in exhaled nitric oxide levels during the Olympic Games (Visit 5 period) compared to the Pre-Olympic periods (Visits 1-4)

Study and Location	Description of Natural Experiment	Chain of Accountability Question	Study Objective(s)	Main Findings	Major Strengths	Major Limitations
			site 650 m from the study school location) during the games compared to before and after the games			
Su et al, 2015 Beijing, China	2008 Summer Olympics – Driving restrictions, industry and construction shutdowns, drastic reductions in all emissions	Did a regulatory action lead to improved air quality and a beneficial human health response?	• Time series analysis of cardiovascular mortality in Beijing residents •Cardiovascular mortality data obtained from Beijing Center for Disease Control for residents of Beijing • Particle number size distribution measurements made on Peking University campus • NO_2, PM_{10} data collected from 8 Beijing monitors • $PM_{2.5}$ measured at 1 location ~5km from Peking University site	• 15.9% to 54.1% reductions in all air pollutant concentrations during games compared to before the games, with increases after the games • 8.8% decreased in cardiovascular mortality associated with interquartile range increases in 1 and 5 day average ultrafine particle counts	• Use of already collected mortality and NO_2 and PM_{10} data • A-B-A design to control for time trends • Coupled with other Beijing Olympic studies investigating mechanistic biomarkers, this study's assessment of cardiovascular mortality provides information on important clinical outcomes, thereby presenting a more complete assessment of health impacts of the air pollution reductions during the Beijing Olympics in adult Beijing residents	• Potential for bias in that Beijing residents may have left Beijing during the games, and thus not be included in a daily count of cardiovascular deaths used in the study. However, authors argue against this since there were no public holidays during the games, and thus little chance to leave the city during the games.
Li et al, 2010 Beijing, China	2008 Summer Olympics – Driving restrictions, industry and construction shutdowns, drastic reductions in emissions from	Did a regulatory action lead to a beneficial human health response?	• Time series analysis of daily counts of outpatient asthma visits for adult residents of urban areas of Beijing • $PM_{2.5}$, O_3, SO_2, NO_2, and	• Number of outpatient visits for asthma were lower (7.3/day) during the games compared to the baseline period (June 2008; 12.5/day) • $PM_{2.5}$ substantially lower (35% to 41%) during the games (46.7 $\mu g/m^3$) compared to the baseline period (78.8	• Use of already collected pollutant and asthma outpatient visit data • Coupled with other Beijing Olympic studies investigating mechanistic biomarkers, this study's assessment of asthma visits	• A-B design may result in residual confounding by time trends and season • Potential for bias in that Beijing residents may have left Beijing during the games, and thus not be

Study and Location	Description of Natural Experiment	Chain of Accountability Question	Study Objective(s)	Main Findings	Major Strengths	Major Limitations
	numerous sources		CO measurements at 3 stations in urban Beijing areas	μg/m^3) and pre-Olympic period (72.3 μg/m^3) • Increases in $PM_{2.5}$ (10 μg/m^3) and O_3 (10 ppb) concentration associated with 2.0% to 4.4% increases in asthma visits	provides an assessment of important clinical outcomes, thereby presenting a more complete assessment of health impacts of the air pollution reductions during the Beijing Olympics in adult Beijing residents	included in a daily count of asthma outpatient visits used in the study
He et al, 2016 Beijing, China	2008 Summer Olympics – Driving restrictions, industry and construction shutdowns, drastic reductions in emissions from numerous sources	Did a regulatory action lead to a beneficial human health response?	• To examine association between increases in monthly concentrations of PM_{10} and monthly mortality rates from January 2006 to December 2010	• Each 10 μg/m^3 reduction in monthly PM_{10} concentration associated with 8.36% to 9.61% reduction in monthly all-cause mortality, and a 8.78% reduction in monthly cardiovascular or respiratory mortality	• Use of large, existing dataset of mortality in Beijing residents • Compared changes in monthly PM_{10} concentration and mortality rates between cities experiencing large reductions in PM_{10} (treated cities) and those with little to no reduction (control cities)	• Although plots suggest there was a greater proportional reduction in PM_{10} concentrations and monthly mortality counts in the cities with large reductions in PM_{10} compared to cities with small concentration reductions, there was not a multivariable analysis to rule out confounding by other temporal and spatial characteristics
Rich et al, 2015 Beijing, China	2008 Summer Olympics – Driving restrictions, industry and construction shutdowns, drastic reductions in emissions from numerous sources	Did a regulatory action lead to a beneficial human health response?	• Cohort study using birth registry data • Pregnant women living in 4 Beijing districts • Same air pollution as Zhang et al (2013)	• Babies whose 8th month of pregnancy was during Olympics were 23 g heavier than babies with their 8th month of pregnancy during the same calendar dates in 2007 or 2009 • No clear pattern for other months of pregnancy • Increases in concentrations of $PM_{2.5}$ (19.8 μg/m^3), NO_2 (13.6 ppb), SO_2 (1.8 ppb), and CO (0.3 ppm) during 8th month associated with	• A-B-A design to control for confounding by time trends • Use of large birth registry to provide ample statistical power • Complimentary analyses of: 1. Changes in birth weight from before to during the games 2. Change in birth weight associated with increased air pollutant concentration	• Potential for residual confounding by women's personal activities during Olympics • No complimentary measurement of biomarkers of potential mechanisms linking air pollution exposure to fetal growth restriction • No control population

Study and Location	Description of Natural Experiment	Chain of Accountability Question	Study Objective(s)	Main Findings	Major Strengths	Major Limitations
				17-34g decreases in birth weight		experiencing same trends in birth weight and fetal growth, but not the air pollution reductions during the Olympics
Lee et al, 2007 Busan, South Korea	2002 Asian Games, 14 days of traffic volume control	Did a regulatory action lead to a beneficial human health response?	Time series analysis of childhood asthma hospitalization rate in year of Asian Games compared to the 3 years before and 1 year after	• 1%-25% reductions in air pollutants in the Post-games period compared to the Pre- and During-Games period • Reduced risk of asthma hospitalizations (RR = 0.73) in children in the "Post-Games" period compared to the "Pre- and During-Games" period • Increases in concentrations of NO_2 (15.09 ppb), SO_2 (3.64 ppb), PM_{10} (28.82 $\mu g/m^3$) and O_3 (12.39 ppb) associated with 24% to 35% increases in the risk of hospital admissions for asthma	• Complimentary analyses of: 1. Changes in a health outcome across 2 periods 2. Change in a health outcome associated with increased air pollutant concentration • A-B-A design comparing year of Asian Games (2002) to years before and after to control for long term time trends	• 14 day period may not be long enough to compare "During-Games" period to "Pre-Games" period as in Beijing Olympic studies • No control population to know if similar patterns of hospitalization and air pollutant concentrations were observed elsewhere