ANNUAL REVIEWS

*Annual Review of Public Health*

# Causal Modeling in Environmental Health

Marie-Abèle Bind

Department of Statistics, Faculty of Arts and Sciences, Harvard University, Cambridge, Massachusetts 02138, USA; email: ma.bind@mail.harvard.edu

ANNUAL REVIEWS CONNECT

www.annualreviews.org

- Download figures
- Navigate cited references
- Keyword search
- Explore related articles
- Share via email or social media

Annu. Rev. Public Health 2019. 40:23–43

First published as a Review in Advance on January 11, 2019

The *Annual Review of Public Health* is online at publhealth.annualreviews.org

https://doi.org/10.1146/annurev-publhealth-040218-044048

Copyright © 2019 by Annual Reviews. All rights reserved

This article is part of a symposium on Causal Inference and Public Health. For a list of other articles in this symposium, see **http://www.annualreviews.org/toc/publhealth/40/1**.

## Keywords

causal inference, public health, epidemiology, air pollution, statistical methods

## Abstract

The field of environmental health has been dominated by modeling associations, especially by regressing an observed outcome on a linear or nonlinear function of observed covariates. Readers interested in advances in policies for improving environmental health are, however, expecting to be informed about health effects resulting from, or more explicitly caused by, environmental exposures. The quantification of health impacts resulting from the removal of environmental exposures involves causal statements. Therefore, when possible, causal inference frameworks should be considered for analyzing the effects of environmental exposures on health outcomes.

## 1. INTRODUCTION

Estimating causal effects in the scientific field of environmental health is often difficult owing to complications related to human subjects. One obvious reason is that it is often unethical to randomize humans to possibly harmful environmental exposures. In many cases, it is also not generally feasible to conduct multifactorial randomized clinical trials. Therefore, to answer causal questions, epidemiologists relied for decades on retrospective data from case-control studies. Epidemiologists now typically collect time series data, manage prospective cohort studies that follow participants over time, or conduct controlled indoor air experiments. Environmental scientists have departed from studying detrimental exposures but instead have conducted creative human experiments randomizing potentially beneficial exposures (e.g., physical activity, green and blue spaces).

Because environmental exposures are nonrandomized in observational studies, there is no guarantee of the environmental exposure to be unconfounded with the (potential) health outcomes. Several strategies have been developed to counter the issue of confounding in observational studies. Of course, if the randomization is not successful at balancing background characteristics across exposure groups, even a randomized clinical trial suffers from the treatment assignment mechanism being confounded with the potential outcomes.

Measures of association between environmental exposures and human health outcomes have been estimated by environmental epidemiology studies (44, 49, 77, 105, 111), which consequently often conclude that further toxicological and/or epidemiological studies are needed to address causality. Whether these environmental health associations in humans are spurious or causal has been discussed in the past (21, 82, 107). To address causality between an environmental exposure and a health outcome, toxicologists have conducted animal studies (71, 104, 109). In contrast, environmental epidemiologists have chosen to quantify the causal effects of environmental exposures on health outcomes in observational human studies using causal inference methods that attempt to tackle the issue of confounding due to the lack of randomized exposures (9, 38, 65). Some authors have focused on reviewing existing causal methods used in environmental health (10, 75). Lewis has discussed the impact of different causal models on estimated effects of disinfection byproducts on preterm birth (66). Other authors have discussed the weight of causal evidence for the associations between an environmental exposure and a health outcome (46, 81, 89, 110).

In Section 2, this article proposes a historical perspective on causal modeling in environmental health with a large focus on air pollution. Section 3 discusses fundamental concepts and potential issues when modeling the observed data with regression. Section 4 presents a nonexhaustive review of the causal modeling methods used in the field of environmental health, with again a focus on air pollution epidemiology. Finally, Section 5 discusses the future of causal modeling in environmental health in the era of big data, machine learning, and greatly improved computational power.

## 2. HISTORICAL PERSPECTIVE: WHY DID SUCCESSFUL STRATEGIES DISAPPEAR?

### 2.1. First Milestones in Environmental Epidemiology

Modern environmental epidemiology is often proposed to have started in the 1850s with the physician John Snow, who was interested in cholera clustering. He identified the public water pump on Broad Street as the source of the cholera outbreak (99). A century later, two major milestones in epidemiology occurred. First, the Great Smog of London, which took place December 5–9, 1952, preceded an unusual increase in mortality in the city, which led to the Clean Air Act of 1956. Another milestone that is often taught in environmental epidemiology is the retrospective British

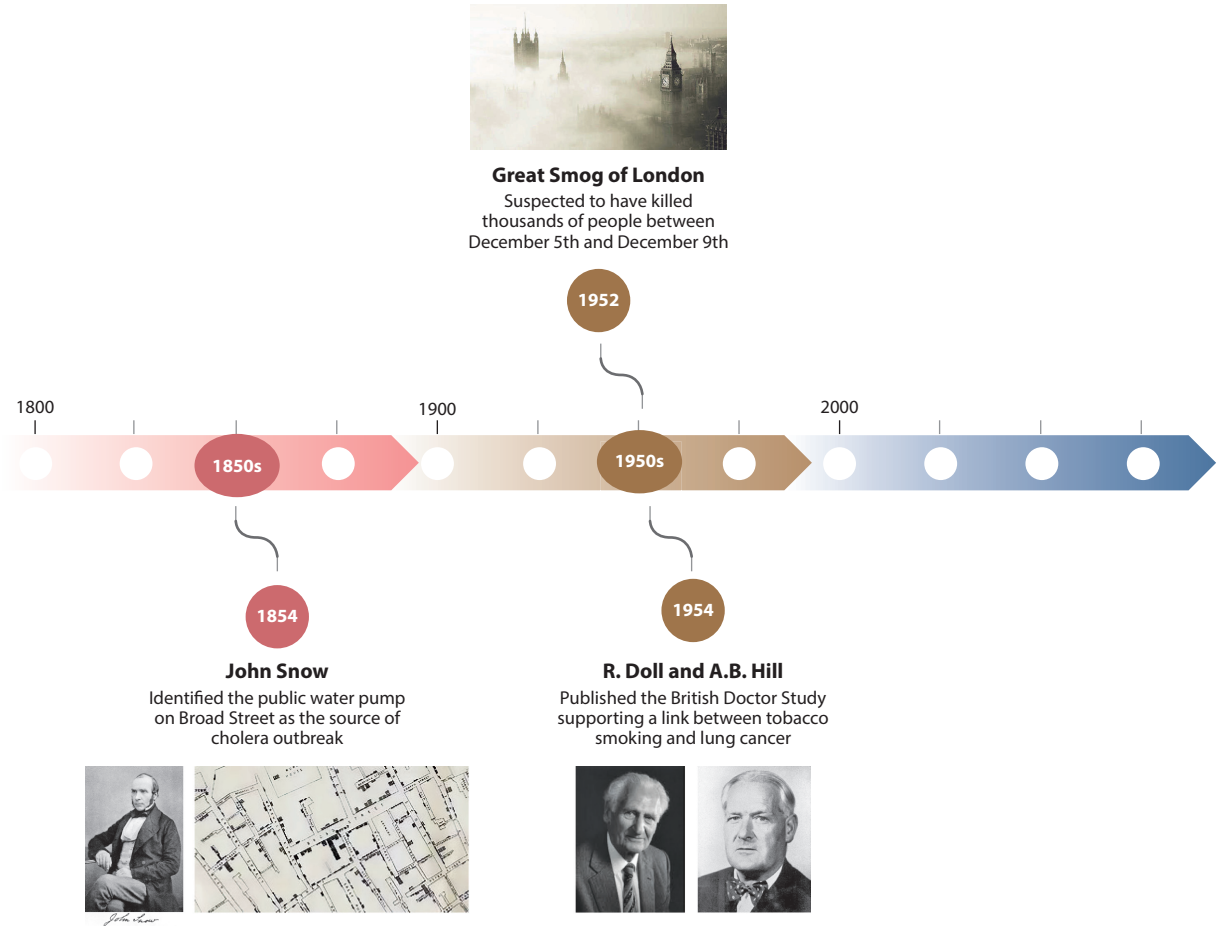


**Figure 1**

Major milestones of environmental epidemiology. Photographs courtesy of Google Images, Wikipedia, The British Library, University of Oxford, and Wikipedia, respectively (*left to right*).

Doctors Study published by Doll & Hill in 1954, which provided some evidence of an association between tobacco smoking and increased risk of lung cancer (32). These major knowledge advances that occurred more than one and a half centuries ago or decades ago, as illustrated in the timeline in **Figure 1**, established modern environmental epidemiology as an attractive field with the potential to discover new medical knowledge using data collected from nonexperimental studies. Even though John Snow, Richard Doll, and Austin Bradford Hill got the right answers, was it the right way to think about and address these causal questions?

## 2.2. Observational Studies and the Matched-Sampling Strategy

Epidemiological studies in biomedicine intend to elucidate whether exposure has some causal effect on health outcomes (17), but the assignment mechanism of these exposures to participants is generally not random and often unknown. Consequently, in such studies, a difference in outcomes between exposed and unexposed groups can be due to confounding by another variable

that is unbalanced in the two groups. In the 1940s, sociologists used matching strategies (12, 47) to limit confounding bias: The goal of matching strategies is to find control units that are similar to exposed units. At that time, powerful computers were not available. Thus, researchers ran into obvious computational difficulties to identify exact pairs of units, for instance one exposed and the other serving as an exposure control. Ideally, the pairing needs to be performed with respect to all confounding variables (101).

In the 1950s, Cochran (15) also suggested using matching to analyze observational studies. Again, this method precedes the outcome analysis in the matched data set comprising the control population similar to the experimental population with respect to disturbing, also known as confounding variables. In the 1960s, he explained how to design, analyze, and summarize observational studies (17). Cochran suggested that researchers design observational studies using matched sampling (e.g., pair matching that consists of finding an approximate twin for each unit) and/or blocking on variables to reduce initial differences between exposed and control groups. He also argued for shifting the focus from estimates with small variance to estimates with small bias (16). Cochran's influence on causal analysis of observational studies was summarized in the 1980s by Rubin (87).

## 2.3. The Prepersonalized Computer Era: Illustrated by Occupational Epidemiology

In the 1970s, epidemiologists were using matched-sampling strategies to handle confounding in cohort studies. For example, Tolenen et al. investigated whether exposure to carbon disulfide was associated with coronary heart disease morbidity and mortality in viscose rayon workers (52, 103). Before comparing (by hand) the means or proportions of health outcomes in workers who had been exposed to rayon versus nonexposed workers, the authors formed cohorts of 343 men and matched these workers with respect to age, district of birth, and similarity of work. Note, these observational studies were not case-control studies, and the matching procedure was appropriately performed with respect to the exposure (and not the outcome).

In the years preceding the 1980s, computing power consisted of centralized management of remote systems; that is, the real computing power and data storage were not on personal desktops. Statistical software created in the 1960s, such as Bio-Medical Data Package (BMDP) at the University of California, Los Angeles, and Statistical Analysis System (SAS) at North Carolina State University, were first operated using batch computing, which explains why, in the prepersonalized computer era, epidemiologists invested a large amount of time thinking about the relevant variables to match on (or to adjust for) before performing statistical analyses. At that time, statistical analyses were tedious tasks, with and even more without, centralized computing power.

The fact that correlation does not imply causation has been a persistent concern in the field of environmental epidemiology. The next section presents the evolution of statistical methods that model conditional associations between air pollution exposures and health outcomes in a nonexhaustive and fairly objective manner. Throughout Section 2.4, the first author's primary affiliation will be flagged if related to the industry for additional context. Let us now comprehend how statistical modeling evolves in the personalized computer era.

## 2.4. The Personalized Computer Era: Illustrated by Air Pollution Epidemiology

Around 1984, there was a shift toward decentralized management of local systems; that is, computing power was available on desktops, and the user managed both software and data storage. Statistical software became increasingly more available on personalized desktops at universities.

The availability of statistical software on personalized computers was a crucial determinant to explain the shift in the 1980s in how researchers conducted statistical analyses in environmental epidemiology. At that time, investigators transitioned away from careful thinking and away from planning the statistical analyses toward easier and faster answers from statistical methods of an observed health outcome regressed on an environmental exposure of interest and background covariates.

In 1984, that is, at the beginning of the personalized computer era, Selvin et al. (96) examined the influence of total suspended particulates, sulfur dioxide, and nitrogen dioxide on mortality across US regions. The authors described fitting 24 linear regressions stratified by sex and adjusted for a set of, what they called, control variables:

$$M_{t,s} \,|\, AP_t,\, C_t \;\; = \beta_{0,s} + \beta_{1,s} AP_t + \beta_{C,s}^T\, C_t + \epsilon_{t,s},$$

where $M_{t,s}$, $AP_t$, and $C_t$ represent the sex-specific daily mortality rate, the daily concentration of air pollutant, and the set of control variables in a particular region, respectively, and $\varepsilon_{t,s} \sim N(0,\, \sigma^2)$. Note, the authors chose a significance level equal to 0.02. Using time series to estimate the conditional association between short-term air pollution and mortality comports some advantages. In a time series setting, the compared units are days, and a major concern is whether days with high air pollution are comparable to days with low pollution and comparable to days with moderate pollution. The distributions of time-varying confounding variables, such as day of the week, month, year, flu, and meteorological variables, are likely to be different across days with low, moderate, and high pollution. However, city-specific distributions of smoking, obesity, medication use, diabetes, age, and sex across days with low, moderate, and high pollution can be assumed to be similar. For example, one can compare mortality counts only between days that are proximate in time.

In 1993, Dockery et al. (31) published the well-known Harvard Six Cities Study, which linked air pollution to mortality using two covariate-adjusted Cox proportional-hazards models stratified by sex and five-year age groups:

$$\lambda_{i,s}(t|City_i,\, C_{it}) = \lambda_{0,s}(t)\;\exp(\beta_{1,1}\mathbb{1}_{i\in City1} + \cdots + \beta_{1,6}\mathbb{1}_{i\in City6} + \beta_C^T C_{it})$$

$$\lambda_{i,s}(t|AP_{City,t},\, C_{it}) = \lambda_{0,s}(t)\;\exp[\beta_{1,1}(\mathbb{1}_{i\in City=1} AP_{City1,t}) + \cdots + \beta_{1,6}(\mathbb{1}_{i\in City6} AP_{City6,t}) + \beta_C^T C_{it}],$$

where $\lambda_{i,s}(t|Z = z)$, $AP_{i\in City,t}$, and $C_{it}$ represent the hazard of dying at time $t$ for individual $i$ and strata $s$ conditional on $Z = z$, the annual city-specific air pollution level where subject $i$ lives, and the subject-specific control variables at time $t$, respectively.

In 1994, Schwartz (92) used a Poisson regression model that controlled for the year of study, continuous time trend of very hot days, temperature, humidity, and winter temperature to estimate the causal relationship between high air pollution exposure and daily mortality (denoted by $M_t$) in Philadelphia:

$$\log[E(M_t)] = \beta_0 + \beta_1 AP_t + \beta_C^T\; C_t \text{ and } E(M_t) = Var(M_t).$$

The same year, Schwartz (91) then used an overdispersed Poisson regression model to associate air pollution exposure with respiratory hospital admissions for the elderly (denoted by $Y_t$) in Detroit,

Michigan:

**Gamble:** investigator from Exxon Biomedical Sciences (in the United States)

$$\log[E(Y_t)] = \beta_0 + \beta_1 AP_t + \beta_C^T\ C_t \text{ and } E(Y_t) = \Phi * Var(Y_t),$$

where $\Phi$ is an overdispersion parameter.

In 1996, Gamble & Lewis (43) reviewed the literature about the associations between air pollution ($PM_{10}$) and health outcomes (e.g., cardiopulmonary morbidity and mortality, hospital admissions for related diseases) and concluded that these associations were weak and that Hill's causal criteria (54) were not met (e.g., temporality, strength of the association, biological plausibility). That year, Dunn & Kingham (35) also suggested that both epidemiological and toxicological evidence are needed to establish a causal link between air quality and health. In 1997, Kuenzli et al. (62) concluded that experimental studies needed to be conducted to investigate the pathophysiological mechanisms involved in the associations between air pollution and health outcomes. In 1998, Bruce et al. (11) concluded (*a*) that intervention studies were required for a strong causal link between indoor air pollution and health outcomes and (*b*) that the quantification of the beneficial effects of the removal of the harmful exposure should be of interest. In 1998, Gamble (42) published another article tackling the plausibility of the $PM_{2.5}$–mortality association, arguing that confounding factors (e.g., physical activity, lung function) could explain the spurious association and that animal studies were not relevant to support an effect in humans because of the different order of magnitudes between experimental conditions versus ambient air pollution levels. In 1999, Schwartz (93) found positive associations between $PM_{10}$ exposure and hospital admissions for heart disease in the elderly in eight US counties using Poisson regression models.

In 2000, Schwartz (93) and Samet et al. (90) attempted to address existent concerns specific to the validity of environmental health studies. The authors introduced a time series setting with copollutants, $AP_{1,t}$, $AP_{2,t}$, and a health outcome, $Y_t$. They introduced a scenario in which $AP_{1,t}$ may confound the $AP_{2,t} - Y_t$ causal relationship:

$$AP_{1,t} = \alpha_0 + \alpha_{1,city} AP_{2,t} + \epsilon_{AP_1,t}, \quad 1.$$

$$Y_t = \beta_0 + \beta_1 AP_{1,t} + \beta_2 AP_{2,t} + \epsilon_{Y,t}, \quad 2.$$

and Equations 1 and 2 are equivalent to

$$Y_t = \beta_0 + \beta_1\alpha_0 + (\beta_2 + \beta_1\ \alpha_{1,city})AP_{2,t} + \beta_1\ \epsilon_{AP_1,t} + \epsilon_{Y,t}. \quad 3.$$

Schwartz and Samet et al. proposed to use two-stage regression, an approach that was popular in social sciences at that time, to assess confounding by the copollutant $AP_{1,t}$. First, in stage 1, $Y_t$ was regressed against $AP_{2,t}$ for each city:

$$Y_t = \delta_0 + \delta_{2,city} AP_{2,t} + \eta_{Y,t}.$$

The estimated $\delta_{2,city}$ has an expectation of $\beta_2 + \beta_1\alpha_{1,city}$. Then, $\alpha_{1,city}$ is estimated using Equation 1. Finally, in stage 2, $\delta_{2,city}$ is regressed against $\alpha_{1,city}$:

$$\delta_{2,city} = \theta_0 + \theta_1\alpha_{1,city} + \kappa_{city}.$$

The two-stage regression approach assumes that $\theta_0$ and $\theta_1$ provide valid estimates of $\beta_2$ and $\beta_1$, respectively. Schwartz and Samet et al. argued that a nonzero estimate of $\beta_1$ suggests a confounded $AP_{2,t}$–$Y_t$ relationship. In 2001, Marcus & Kegler (69) argued against the two-stage regression strategy by proposing counterexamples and simulations. For instance, he introduced a third variable $AP_{3,t}$ such that the $AP_{2,t}$–$Y_t$ relationship is confounded not only by $AP_{1,t}$, but also by $AP_{3,t}$. The authors argued that in that third-confounder setting the two-stage approach would incorrectly estimate the magnitude of the $AP_{2,t}$–$AP_{1,t}$–$Y_t$ pathway.

**L.A. (Tony) Cox:** investigator from Cox Associates (located in the United States)

In 2002, Wong et al. (108) argued that there would be some causal evidence if effect estimates were similar when modeling the association between air pollution and hospital admissions in Hong Kong and London with the same Poisson regression model. The same year, Le Tertre et al. (64) concluded in favor of a causal interpretation of the air pollution–mortality association. Using overdispersed Poisson regression models, the authors estimated air pollution–mortality conditional associations in nine French cities. After conducting a meta-analysis, they obtained a significant and positive pooled estimate, even though the nine estimates were not always significant.

In 2007, Pope & Burnett (80) disagreed with the argument proposed by Janes et al. (58), who, after observing air pollution–mortality associations at the national scale but not at the local scale, asserted that the association between long-term exposure to air pollution and mortality is more likely to be confounded at the national scale than on the local scale. The same year, Brook (10) also proposed his view on the causal question and concluded, on the basis of the existing epidemiological studies, that the final evidence was now sufficient to implicate particulate matter exposure as a cause of cardiovascular disease. Goldberg (45) also argued that standard epidemiological designs could not uniquely identify any individual component of air pollution as a causal agent of a health effect. He also suggested that researchers examine precisely the toxicity of mixtures of air pollutants. He finally proposed toxicological studies and human controlled studies that capitalized on multifactorial designs to study specific component(s) of the air pollution mixture.

In 2009, Knol et al. proposed an elicitation based on fourteen European experts, such as epidemiologists, toxicologists, and clinicians, to examine the likelihood of (*a*) a causal relationship and (*b*) causal pathways between ultrafine particles and key health outcomes. The experts concluded a medium-to-high likelihood for a causal relationship (61). The same year, to understand potential causal mechanisms of the air pollution–cardiovascular relationship, Delfino et al. (27) conducted a panel study on elderly participants with coronary artery disease and investigated whether traffic-related pollutants were associated with inflammatory biomarkers. The authors reported statistically significant positive associations using mixed-effects models (27):

$$Y_{ij} = \beta_0 + u_i + \beta_1 AP_{ij} + \beta_C^T C_{ij} + \epsilon_{ij},$$

where $Y_{ij}$ and $u_i$ represent a health outcome for subject $i$ at visit $j$ and random intercepts, respectively.

In 2012, L.A. (Tony) Cox et al. examined the association between short-term exposure to air pollution and daily mortality in 100 US cities. The authors concluded that the $PM_{2.5}$–mortality relationship was not causal because of nonlinear confounding by temperature when modeled with splines (25). This study was a reanalysis of an earlier epidemiological study that found mortality associations with $PM_{2.5}$ and $PM_{10}$ (33). The conclusions of Cox et al. using Bayesian model averaging conflicted with the conditional associations estimated by Dominici and colleagues (33) using a Bayesian hierarchical model, including, in particular, smooth functions of same-day temperature with six degrees of freedom. The same year, Devlin et al. (28) reported causal effects of ozone on cardiovascular markers in human controlled exposure chambers. Padula et al. (76) also estimated

the causal effects of traffic-related air pollution during pregnancy and birth weight using machine-learning methods. They concluded that the absence of arbitrary model assumptions should result in relatively unbiased estimates (76). In 2013, Cox (21) argued for the possibility of spurious spatial exposure–response associations in environmental studies and against the causal interpretation of linear regression coefficients in air pollution epidemiology (22). The same year, Jarjour et al. (59) conducted a scripted exposure crossover study and found no change in lung function when comparing bikers cycling on low-traffic versus high-traffic routes. In 2014, Hackstadt et al. (50) used principal stratification in a randomized air cleaner intervention trial to quantify the causal effects of the intervention in the subgroup of units for which the air cleaner would have reduced indoor air pollution. In 2015, Cox (20) used time series to assess whether historical air pollution concentrations helped to predict and explain changes in cardiovascular and all-cause mortality, and he concluded that reducing $PM_{2.5}$ and $O_3$ did not result in human longevity benefits. This strategy is often referred to in economics as conducting a Granger causality test (more on this below). However, the authors found no evidence of an association between air pollution and mortality and incorrectly accepted the null hypothesis of no association. That same year, Schwartz et al. (94) estimated positive causal effects of $PM_{2.5}$ on mortality by capitalizing on an instrumental variable approach that used back trajectories of air masses over Boston for a six-year period as random instruments. Back trajectories are the most likely estimated paths of air masses, assuming no dispersion in the atmosphere.

In 2016, after randomized mouse studies had demonstrated that secondhand smoke has some causal effect on long interspersed nuclear element 1 (LINE-1) methylation (104), and after an air pollution–methylation association had been observed in observational studies (5), Chen et al. (14) performed a causal mediation analysis within a randomized crossover trial and estimated a mediated effect of air pollution on cardiovascular markers via a decrease in LINE-1 DNA hypomethylation. A creative epidemiological study reported positive associations between exposure to ambient air pollution and mortality using observational data in cows (19). The authors concluded that these findings reinforce the evidence on the plausibility of the causal health effects of air pollution exposures. Wang et al. (106) also quantified the causal effect of long-term $PM_{2.5}$ exposure on mortality using a difference-in-difference approach. Briefly, this popular method in economics estimates the causal effect of an intervention (e.g., pollution reduction in Boston) on a time series outcome in a population (mortality in Boston at time $t$) by estimating the missing potential outcome of no intervention using, for example, the mortality at time $t$ of a similar city that did not implement the intervention (similar with respect to the outcome time trend).

In 2017, Schwartz et al. (95) used an instrumental variable approach again and estimated positive causal effects of local air pollution ($PM_{2.5}$ and black carbon) on mortality. The authors used Granger causality to assess residual confounding. The combination of the height of the planetary boundary layers and wind speed was used to construct the instrument for $PM_{2.5}$, black carbon, and $NO_2$. The same year, Sheldon & Sankaran (97) performed a two-stage least squares analysis and concluded that Indonesian fires caused an increase in acute respiratory tract infections and acute conjunctivitis. Makar et al. (68) also estimated the causal effect of low-level $PM_{2.5}$ on hospitalization using inverse probability weighting (IPW), a method used in survey sampling. Briefly, IPW used the estimated propensity score to create a hypothetical population in which the exposure assignment can be assumed to be independent of the potential outcomes given the confounding variables.

Cox et al. (23) used machine-learning methods to examine the causal link between short-term exposure to air pollution and cardiovascular-related hospital admissions among older adults. Using random forest and Bayesian network learning, the authors concluded with strong statements that

geographic location, time, and temperature, but not $PM_{2.5}$, caused cardiovascular-related hospital admissions, again accepting the null hypothesis when failing to reject it. Although the authors claimed strong causal conclusions regarding the etiology of cardiovascular-related hospital admissions, they also concluded in their article (23) that better causal modeling methods are needed to better comprehend how reducing air pollution would affect public health.

Bind et al. (8) used quantile mediation analysis and found evidence that the association between air pollution and fibrinogen may be mediated by interferon-$\gamma$ methylation, but only for specific quantiles of the mediator and outcome distributions. In 2018, Fiorito et al. (39) designed a nested case-control study and conducted regression-based mediation analysis in nonsmokers. They concluded that the air pollution associations with cardio- and cerebrovascular disease are mediated by oxidative stress and inflammation. That same year, Zigler et al. (113) also performed a causal analysis to quantify the impact of US ambient air quality standard nonattainment on health outcomes.

In 2018, Cole-Hunter et al. (18) conducted a repeated measures study and quantified associations between air pollution and cardiopulmonary outcomes in healthy adults using multivariate-adjusted linear mixed-effect models. Air pollution exposure was estimated at the participant's residential and occupational addresses. The authors performed mediation analyses and reported that occupational-address noise and residential greenness (but not physical activity) mediated the air pollution–cardiovascular associations. It is trendy now to conduct scripted exposure studies to estimate causal relationship in environmental epidemiology. Recently, Sinharay et al. (98) conducted a crossover study randomizing elderly participants to a two-hour walk either along Oxford Street in London or in an urban park. The authors reported that walking in areas with high levels of traffic pollution induced adverse cardiopulmonary effects (98).

## 2.5. The Analysis of Comparable Groups Makes a Comeback

Balancing strategies, such as matching (17), can help create similar groups with respect to important background variables. The construction of comparable groups is now facilitated by modern computing power and available software (51, 117). In the past few years, statistical analysis of comparable groups has become popular again in environmental epidemiology. In 2012, Moore et al. (72) illustrated the use of causal models for realistic individualized exposure rules (CMRIER) that generalize marginal structural models to study the effect of ozone on asthma-related hospital discharge. In 2015, Schwartz et al. (94) estimated a positive causal effect of air pollution on mortality using a quasi-Poisson model within deciles of the estimated propensity scores. In 2017, Baccini et al. (1) assessed the short-term impact of $PM_{10}$ (dichotomized at 40 $\mu g/m^3$) on the total number of attributable deaths in Milan using propensity score matching. The same year, Bind & Rubin (6) argued that the obtainment of comparable groups is a necessary step before comparing health outcomes between exposed and unexposed. Before comparing children's lung function, the authors illustrated how to use matching strategies to create two similar groups of children (similar with respect to age, sex, height): one composed of children with smoking parents and the other composed of children with nonsmoking parents. The effort to compare apples to apples is crucial in nonrandomized settings to depart from association and to address causation (100).

The achievement of covariate balance should be presented before estimating causal effects in both randomized and nonrandomized studies to convince journal readers and policy makers (8, 86). Transparent balance checks allow constructive discussions. If a reader disagrees with the author's conclusion, for instance, because a variable was omitted from the statistical analysis, the opportunity exists to consider that particular variable in the next analysis and to present the subsequent results at the next debate.

## 2.6. Centralized Management of Remote Systems Also Makes a Comeback

History also repeats itself in terms of computing management. Some investigators have departed from decentralized computing of local systems to return to centralized management of remote systems. With the recent availability or collection of big data in environmental health, researchers need an adapted computing platform, such as cloud computing, to harness the burden of high-dimensional statistics analyses in biomedical research. Griebel et al. (48) review the recent increase in cloud computing in health care studies.

# 3. REGRESSION: WHAT ARE YOU IMPLICITLY IMPUTING?

## 3.1. The Possible Danger of the Implied Imputation

Most environmental studies have directly modeled the observed outcome eschewing, for instance, the definition of a causal estimand as a function of the potential outcomes, the comparison of comparable groups, and the consideration of a joint model for the potential outcomes. This common analysis strategy can be problematic when obtaining valid causal conclusions.

**3.1.1. Causal inference: a missing data problem.** Causal inference is inherently a missing data problem (55). Therefore, researchers interested in quantifying causal effects must somehow impute the missing potential outcome for each unit. Modeling the observed outcome as a function of covariates and Gaussian noise can be erroneous because it misses the essential concept of causal inference (i.e., being a missing data problem). Interesting papers have showed differences in estimates of causal effects in settings where (*a*) the exposure was randomized (gold standard setting); (*b*) the exposure was not randomized and a regression was fitted using the observed exposure, covariates, and outcome (63); and (*c*) a data set was reconstructed so that the exposure could be assumed to be randomized given covariates (26). Not surprisingly, the estimated causal effects of the first and third strategies were close. The third strategy has an implicit imputation of the missing potential outcomes capitalizing on the pairing of matched units.

**3.1.2. The need for defining a causal estimand, a conceptual stage, and a design stage.** Environmental epidemiologists need to ponder and lock several considerations before attempting to address causal questions: for example, How would I proceed if I could conduct a randomized experiment in the population of interest? What is the causal estimand? How can the nonrandomized data be described as collected from a randomized experiment? How can I use subject matter knowledge to choose the variables to balance? Which design strategy will be performed? Which statistical method should be used to compare the balanced exposed and unexposed groups?

Posing the causal question and defining the causal estimands as functions of the potential outcomes are needed (86). In 2014, Zigler & Dominici (114) argued for well-defined actions in terms of potential outcomes in air pollution epidemiology. In 2016, Zigler et al. (116) argued for the use of the potential outcome framework to examine the causal effect of air quality regulations on long-term health effects.

It is evident that a conceptual stage in environmental epidemiology is also needed. In 2017, Bind & Rubin (6) advocated for a multistage strategy (i.e., with conceptual, design, statistical, and summary stages) to emphasize how critical the conceptual stage is to draw statistical conclusions. In particular, the authors argued that *p*-values originating from analysis of observational studies have no scientific validity without embedding the observational study into a plausible hypothetical randomized experiment; that is, epidemiologists should state a plausible randomization mechanism before calculating a *p*-value.

Rubin (88) has promoted for decades a separation of the design stage from the analysis stage in observational studies and illustrated his strategy in the context of the US tobacco litigation. Using propensity score matching, he created two comparable groups of male current smokers and of male never smokers, blindly of any outcome data (88). Dominici & Zigler (34) have also discussed how to evaluate the evidence of causality in air pollution epidemiology. The authors proposed criteria to evaluate evidence of causality in environmental epidemiology: (*a*) What actions or exposure levels are being compared? (*b*) Was an adequate comparison group constructed? and (*c*) How closely do these design decisions approximate an idealized randomized study? Each of these three points can be thought as arguing for the need to, respectively, define a causal question in terms of potential outcomes, have a successful design stage, and have a plausible conceptual stage.

### 3.2. Causal Estimands

Researchers typically use standard regression to model the mean of the observed outcome, often restricting themselves to the selection of a causal estimand that is not necessarily of scientific interest (e.g., average causal effect). In environmental epidemiology, some evidence has shown that the associations can be stronger at the tail of the observed outcome distribution (4). Instead, scientists that impute the missing potential outcomes can choose causal estimands of interest because any function of the potential outcomes (that have been imputed) can then be calculated.

### 3.3. Finite versus Super-Population Interpretation and Generalizability

The frequentist approach assumes that the units consist of a random sample from an infinite population, whereas the joint Bayesian model of [Y(0),Y(1)] assumes that randomness comes from the exposure assignment and therefore does not extrapolate its conclusions to a population beyond the finite sample. Imbens & Rubin (57) discussed technical differences between (*a*) the frequentist linear regression using the observed outcome and (*b*) the Bayesian model of the joint distribution of the potential outcomes [Y(0),Y(1)]. Regression coefficients and their confidence intervals estimated by the former approach have a superpopulation interpretation, and in contrast, the posterior mode and credible interval obtained by the latter strategy have a finite population interpretation. They also compared the assumptions of both approaches. For example, the Bayesian model can be more flexible because it does not automatically constrain the residual variance to be the same in the exposed and unexposed groups.

The difference in interpretation between finite and superpopulation leads inherently to the topic of generalizability. Although randomized clinical trials are optimal to address causality in humans, study populations of observational studies are generally chosen such that they can be generalizable to an entire population of interest. Besides, observational studies are often the only ethical and feasible design to study health effects of multiple environmental stressors or pollutants in humans. Consequently, causal modeling has an important role to play in environmental epidemiology.

## 4. CAUSAL MODELING METHODS TARGETING ENVIRONMENTAL HEALTH SCIENTISTS

### 4.1. Experimental Studies

Even though randomized experiments are the gold standard for addressing causality, in practice they can also suffer from covariate imbalance if the randomization was not successful. To avoid this

issue, Morgan & Rubin (73) proposed a rerandomization strategy to improve covariate balance in experiments. Air pollution chamber studies have been conducted in humans and analyzed with mixed-effects models and with paired *t*-tests that assume Gaussian asymptotic distributions (28, 29, 112). In randomized experiments, an actual randomization occurred and should be incorporated into the statistical analysis. Combining Bayesian and randomization-based inferences, as suggested by Bind & Rubin (6), could also be useful in human experimental settings with a small number of units. Although expertise in causal inference is important when conducting experimental studies, mastering classical insights of experimental design and essential concepts in causal inference is particularly crucial when conducting observational studies in which the environmental exposure mechanism is unknown and nonrandomized.

## 4.2. Estimating Main Causal Effects in Observational Studies

We present the most common methodological strategies that have been used to estimate the main effects of an environmental exposure on a health outcome. Some epidemiological studies attempted to tackle the confounding issue capitalizing on the fact that "association is causation in the absence of confounding"; others have attempted to obtain a setting in which plausible imputations of the missing potential outcomes can be performed. We now present the most common approaches that have been used to estimate the main effects of an environmental exposure on a health outcome.

**4.2.1. Directed acyclic graphs.** In 2011, Flanders et al. (40) constructed an indicator that detected residual confounding in nonrandomized studies based on directed acyclic graphs (DAGs) and capitalized on regression coefficients estimated with a future exposure variable. In 2012, before obtaining an association between $NO_2$ exposure and increased risk of ischemic heart disease, Beckerman et al. (3) constructed a fairly comprehensive DAG to identify potential confounders, such as socioeconomic status (see **Figure 2**). In 2015, Weisskopf et al. (107) used DAGs to argue

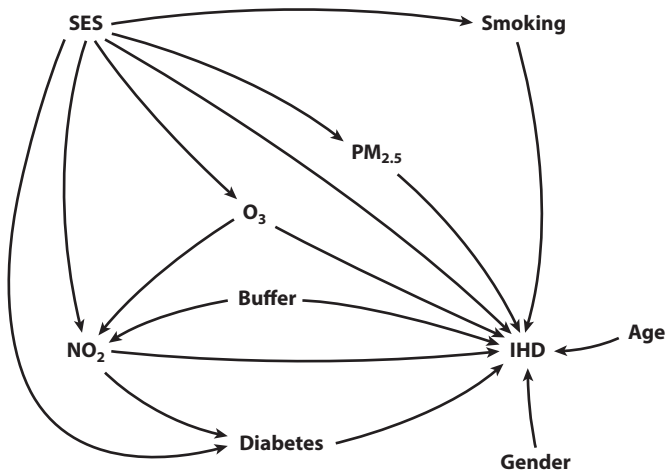


**Figure 2**

Example of a directed acyclic graph (DAG) in environmental epidemiology. Figure adapted from Beckerman et al. (3). Abbreviations: IHD, ischemic heart disease; $PM_{2.5}$, particulate matter with a diameter of 2.5 μm or less; SES, socioeconomic status.

that the air pollution–autism association was not confounded by time-invariant factors and concluded that the causal association between air pollution and autism was increasingly compelling.

**4.2.2. Propensity score.** Because the exposure mechanism is unknown in observational studies, researchers have attempted to estimate the distribution of the exposure. The propensity score is the probability of a unit to be assigned to a particular exposure level given a set of covariates. Matching units with respect to their estimated propensity score has been a successful strategy to obtain balanced groups (88). In 2017, Baccini et al. (1) argued for this approach to address causality in environmental epidemiology because it does not require any modeling of the confounder–outcome relationship nor does it require model extrapolation.

**4.2.3. g-methods.** Instead of modeling the observed exposure, covariates, and outcome, the g-estimation method models the potential outcomes under no exposure as a function of observed exposure, covariates, and outcome (79, 84). Using this method, Moore et al. obtained a nonsignificant estimate of the effect of ozone reductions on the proportion of asthma-related hospital discharges (72).

**4.2.4. Mendelian randomization.** Mendelian randomization relies on the principle that if a genetic variant has an effect on an environmentally modifiable exposure that itself has an effect on a disease, then this genetic variant should also be related to the disease. Relton & Davey Smith (83) suggest that there is a potential for several applications in environmental epidemiology because DNA methylation has been associated with environmental exposures such as smoking, drinking, arsenic, and ambient air pollution.

## 4.3. Methods Adapted from the Field of Economics

Environmental health studies that reported causal main effects also used causal methods that emerged in economics.

**4.3.1. Granger causality.** In contrast with its name, the Granger causality approach is not a proper causal inference method in the sense that it focuses on prediction and does not focus on the imputation of the missing potential outcomes. The approach consists of testing whether one time series can be used to predict another. In 2015, Farhani & Ozturk (37) were interested in testing the environmental Kuznets curve (EKC) hypothesis, which postulates an inverted-U-shaped relationship between different pollutants and per capita income. The authors used an auto-regressive distributed lag approach with Granger causality models to examine the causal relationship between $CO_2$ emissions, real gross domestic product, energy consumption, financial development, trade openness, and urbanization in Tunisia (37). In 2018, Chen et al. (13) used a modified Granger causality test and generalized autoregressive conditional heteroscedastic (GARCH) models to examine the causal relationship between ambient fine particles and human influenza in Taiwan. In 2018, Jiang & Bai (60) conducted a Granger causality test to assess whether the air pollution concentration of a given Chinese city was affected by air pollution from neighboring cities.

**4.3.2. Difference in differences.** As mentioned in Section 2.4, difference-in-difference estimators can be used to estimate causal effects on mortality at time $t > T_0$ in the

population of Boston of a pollution reduction intervention that occurred at time $T_0$ defined as $\tau_{Boston,t} = E_{i \in Boston}[M_{i,t}(\text{intervention} = 1) - M_{i,t}(\text{intervention} = 0)]$, for $t > T_0$. Obviously, $M_{i,t}(\text{intervention} = 0)$ is a nonmeasurable potential outcome but can be estimated using the mortality of another similar city at time $t$ (similar with respect to outcome time trend). Wang et al. (106) applied this approach and estimated the causal effects of long-term exposure to $PM_{2.5}$ on mortality in New Jersey using covariate-adjusted regression models.

**4.3.3. Regression discontinuity.** The regression discontinuity approach consists of studying the effect of an exposure (dichotomized at a certain threshold) on an outcome; that is, two groups of units are defined by their exposure level being either below or above a threshold. Units that are close to the threshold can be assumed to be very similar and as if they would have been randomized to one side or the other (102). This approach capitalizes on this approximated randomization to compare outcomes between units that are close to the threshold but on either side. In 2017, Ebenstein et al. (36) used the regression discontinuity design to estimate the effect of $PM_{10}$ on life expectancy in China using the Huai River as the threshold, also known as the boundary variable.

## 4.4. Intermediate Variables

Variables in the causal pathway between the exposure and outcome variables are often called intermediate or mediator variables. In practice, it is advised not to directly adjust for intermediate variables in regression analysis to estimate unbiased main effects. However, the mediated effect has been of interest in environmental health.

**4.4.1. Principal stratification.** Methods using instrumental variable are particular cases of principal stratification, which considers an intermediate variable (in the sense of posttreatment variable) in the statistical analysis (41). Zigler et al. (115) used principal stratification to "examine causal effects of a regulation on health that are and are not associated with causal effects of the regulation on air quality." Hackstadt et al. (50) also used principal stratification in a randomized air cleaner intervention trial to examine the extent to which an effect of an environmental intervention on health outcomes coincides with its effect on indoor air pollution.

**4.4.2. Mediation.** In 2016, Bind et al. (7) derived mediation formulae for examining the mediated effect of exogenous exposures, such as air pollution. The authors illustrated the method by estimating the mediated effect of two environmental exposures, air pollution and temperature, on the intercellular adhesion molecule 1 (ICAM-1) via a change in *ICAM-1* DNA methylation. In 2017, Balte et al. (2) conducted a path analysis using linear mixed-effects models to examine the causal pathways among prenatal and childhood factors that impact lung function later. Path analysis is used to describe dependencies among variables and can be viewed as a special case of structural equation modeling. In 2018, Cole-Hunter et al. (18) conducted a regression-based mediation analysis to examine occupational-address noise, residential-neighborhood greenness, and total daily physical activity as possible mediators in the association between $PM_{10}$ exposure and cardiopulmonary outcomes. The same year, Hüls et al. (56) also conducted a regression-based mediation analysis and concluded that the association between air pollution and impairment in visuoconstruction performance (a measure of cognitive function) was mediated by lung function.

## 5. THE FUTURE OF CAUSAL MODELING IN ENVIRONMENTAL HEALTH

### 5.1. Examples of Outcomes Associated with Environmental Exposures for Which Epidemiologists Need to Address Causality

In 2018, in the context of environmental exposure effects on autism, Hertz-Picciotto et al. (53) discussed whether evidence from observational associations permits causal inference. Also in 2018, Lu et al. (67) associated air pollution with criminal and unethical behaviors using a 9-year panel of more than 9,000 US cities and conducted a psychological experiment to validate their findings. Examining causal links between air pollution exposures and complex and still poorly understood health outcomes, such as autism, multiple sclerosis, Parkinson's disease, and types of dementia may be the next challenge for environmental epidemiologists.

### 5.2. Big Data

In 2012, Padula et al. (76) quantified the causal effects of exposure to traffic-related air pollution on birth weight using machine learning and targeted maximum likelihood estimation (TMLE), which, instead of minimizing variance or mean square errors, targets the maximum likelihood estimate but in a way that reduces bias. In 2013, Diaz & van der Laan (30) proposed to use inverse probability of treatment weighted (IPTW), augmented IPTW (a doubly robust estimator), TLME, and stochastic interventions to assess causality. In 2014, Mauderly and coauthors (70, 71) performed a multiple additive regression tree analysis in a multipollutant air quality study in rodents. In 2017, Oulhote et al. (74) combined ensemble learning (i.e., SuperLearner) and g-computation to estimate the effect of chemical mixtures. In 2018, Golan et al. (44) discussed the big data paradigm. In their paper, they performed regression-based analyses and provided association estimates between environmental exposures and fetal growth. In 2017, Cox et al. used (*a*) nonparametric models to avoid model misspecification and (*b*) ensemble modeling that averaged over several model results. This strategy has become popular because it can yield estimates with lower bias and variance (24). Comprehending the statistical properties of black box machine-learning methods is currently an area of intense research and should lead to a better understanding of which machine-learning methods to use in various causal inference settings.

### 5.3. Computational Power

Computers have become increasingly more powerful, leading to the possibility of returning to exact causal inference instead of relying on asymptotics and modeling (78). As stated before, combining Bayesian modeling with Fisherian exact inference (8) is now possible by harnessing greatly improved computational power. Monte Carlo methods can now be used to compute in parallel posterior distributions of more complex causal estimands.

Because of the emergence of approximate solvers in the past decade, existing solutions for the old Monge–Kantorovich mass transportation problem can now be computed efficiently for large data sets (78). In special cases, the Monge–Kantorovich problem is equivalent to the optimal matching problem. In 2012, Rosembaum proposed to perform combinatorial optimization and to use linear programming methods for optimal matching in causal inference (85). Software is now available to perform the task (51, 117). Focusing on optimal balanced data sets may not be as optimal for addressing causality in observational studies, and the statistical analyses of combined nonoptimal matched data sets may provide better statistical properties such as coverage and efficiency.

## 6. CONCLUSION: WHY SHOULD YOU WORRY ABOUT CAUSATION VERSUS ASSOCIATION?

Important concepts of experimental design and causal inference are needed for readers of environmental epidemiology to evaluate the quality of papers and for epidemiologists to address causal questions with success. Because of the lack of randomization, significant results from observational studies should always be interpreted and evaluated with caution. In practice, the ignorability assumption is often strong and unrealistic. Sensitivity analyses, such as tipping point analyses, should be performed. Reproducibility of the findings, especially from observational studies, is also essential for causal inference modeling to draw conclusions, to intervene, and to change policies. In addition to learning regression models, environmental scientists should be trained in experimental and quasi-experimental designs, as well as essential concepts of causal inference. Although there is a need to address causality in environmental epidemiology, additional issues should be tackled. For example, before considering causal modeling, environmental scientists should address the issues of missing data (e.g., perform multiple imputation of the missing values, sensitivity analyses such as tipping point analyses), measurement error, model selection, loss to follow-up, time-varying confounding, multiple testing, and reproducibility. Researchers should also consider only models that were predetermined in protocols avoiding the issue of *p*-value hacking during the model selection phase. For decades, environmental epidemiology has focused on observational designs, and therefore their findings suffered from the fact that association is not causation. Implementing the multistage causal strategy coupled with interdisciplinary collaborations will be crucial to address the next century's challenges in environmental epidemiology.

## DISCLOSURE STATEMENT

The author is not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review. The content is solely the responsibility of the author and does not necessarily represent the official views of the National Institutes of Health.

## ACKNOWLEDGMENTS

The author thanks Francesco Forastiere, for providing historical context of epidemiological practice in the 1970s, and the reviewer, for insightful comments during the revision process. Research reported in this publication was supported by the John Harvard Distinguished Science Fellows Program within the Faculty of Arts and Sciences (FAS) Division of Science of Harvard University and by the Office of the Director, National Institutes of Health under award DP5OD021412.

## LITERATURE CITED

1. Baccini M, Mattei A, Mealli F, Bertazzi PA, Carugno M. 2017. Assessing the short term impact of air pollution on mortality: a matching approach. *Environ. Health* 16:7
2. Balte P, Kühr J, Kruse H, Karmaus W. 2017. Body burden of dichlorodiphenyl dichloroethene (DDE) and childhood pulmonary function. *Int. J. Environ. Res. Public Health* 14(11):1376
3. Beckerman BS, Jerrett M, Finkelstein M, Kanaroglou P, Brook JR, et al. 2012. The association between chronic exposure to traffic-related air pollution and ischemic heart disease. *J. Toxicol. Environ. Health A* 75:402–11
4. Bind MA, Coull BA, Peters A, Baccarelli AA, Tarantini L, et al. 2015. Beyond the mean: quantile regression to explore the association of air pollution with gene-specific methylation in the Normative Aging Study. *Environ. Health Perspect.* 123:759–65

5. Bind MA, Lepeule J, Zanobetti A, Gasparrini A, Baccarelli A, et al. 2014. Air pollution and gene-specific methylation in the Normative Aging Study: association, effect modification, and mediation analysis. *Epigenetics* 9:448–58
6. Bind MA, Rubin DB. 2017. Bridging observational studies and randomized experiments by embedding the former in the latter. *Stat. Methods Med. Res.* 962280217740609
7. Bind MA, Vanderweele TJ, Coull BA, Schwartz JD. 2016. Causal mediation analysis for longitudinal data with exogenous exposure. *Biostatistics* 17:122–34
8. Bind MA, VanderWeele TJ, Schwartz JD, Coull BA. 2017. Quantile causal mediation analysis allowing longitudinal data. *Stat. Med.* 36:4182–95
9. Brauer C, Budtz-Jørgensen E, Mikkelsen S. 2008. Structural equation analysis of the causal relationship between health and perceived indoor environment. *Int. Arch. Occup. Environ. Health* 81:769–76
10. Brook RD. 2007. Is air pollution a cause of cardiovascular disease? Updated review and controversies. *Rev. Environ. Health* 22:115–37
11. Bruce N, Neufeld L, Boy E, West C. 1998. Indoor biofuel air pollution and respiratory health: the role of confounding factors among women in highland Guatemala. *Int. J. Epidemiol.* 27:454–58
12. Chapin FS. 1947. *Experimental designs in sociological research*. New York: Harper
13. Chen CWS, Hsieh YH, Su HC, Wu JJ. 2018. Causality test of ambient fine particles and human influenza in Taiwan: age group-specific disparity and geographic heterogeneity. *Environ. Int.* 111:354–61
14. Chen R, Meng X, Zhao A, Wang C, Yang C, et al. 2016. DNA hypomethylation and its mediation in the effects of fine particulate air pollution on cardiovascular biomarkers: a randomized crossover trial. *Environ. Int.* 94:614–19
15. Cochran WG. 1953. Matching in analytical studies. *Am. J. Public Health Nations Health* 43:684–91
16. Cochran WG. 1968. The effectiveness of adjustment by subclassification in removing bias in observational studies. *Biometrics* 24:295–313
17. Cochran WG, Chambers SP. 1965. The planning of observational studies of human populations. *J. R. Stat. Soc. Ser. A* 128:234–66
18. Cole-Hunter T, de Nazelle A, Donaire-Gonzalez D, Kubesch N, Carrasco-Turigas G, et al. 2018. Estimated effects of air pollution and space-time-activity on cardiopulmonary outcomes in healthy adults: a repeated measures study. *Environ. Int.* 111:247–59
19. Cox B, Gasparrini A, Catry B, Fierens F, Vangronsveld J, Nawrot TS. 2016. Ambient air pollution-related mortality in dairy cattle: Does it corroborate human findings? *Epidemiology* 27:779–86
20. Cox LA Jr., Popken DA. 2015. Has reducing fine particulate matter and ozone caused reduced mortality rates in the United States? *Ann. Epidemiol.* 25:162–73
21. Cox LA Jr., Popken DA, Berman DW. 2013. Causal versus spurious spatial exposure-response associations in health risk analysis. *Crit. Rev. Toxicol.* 43(Suppl. 1):26–38
22. Cox LAT Jr. 2013. Caveats for causal interpretations of linear regression coefficients for fine particulate (PM2.5) air pollution health effects. *Risk Anal.* 33:2111–25
23. Cox LAT Jr. 2017. Do causal concentration-response functions exist? A critical review of associational and causal relations between fine particulate matter and mortality. *Crit. Rev. Toxicol.* 47:603–31
24. Cox LAT Jr., Liu X, Shi L, Zu K, Goodman J. 2017. Applying nonparametric methods to analyses of short-term fine particulate matter exposure and hospital admissions for cardiovascular diseases among older adults. *Int. J. Environ. Res. Public Health* 14:1051
25. Cox T, Popken D, Ricci PF. 2012. Temperature, not fine particulate matter (pm2.5), is causally associated with short-term acute daily mortality rates: results from one hundred United States cities. *Dose Response* 11:319–43
26. Dehejia RH, Wahba S. 1999. Causal effects in nonexperimental studies: reevaluating the evaluation of training programs. *J. Am. Stat. Assoc.* 94:1053–62
27. Delfino RJ, Staimer N, Tjoa T, Gillen DL, Polidori A, et al. 2009. Air pollution exposures and circulating biomarkers of effect in a susceptible population: clues to potential causal component mixtures and mechanisms. *Environ. Health Perspect.* 117:1232–38
28. Devlin RB, Duncan KE, Jardim M, Schmitt MT, Rappold AG, Diaz-Sanchez D. 2012. Controlled exposure of healthy young volunteers to ozone causes cardiovascular effects. *Circulation* 126:104–11

29. Devlin RB, Smith CB, Schmitt MT, Rappold AG, Hinderliter A, et al. 2014. Controlled exposure of humans with metabolic syndrome to concentrated ultrafine ambient particulate matter causes cardiovascular effects. *Toxicol. Sci.* 140:61–72
30. Diaz I, van der Laan MJ. 2013. Assessing the causal effect of policies: an example using stochastic interventions. *Int. J. Biostat.* 9:161–74
31. Dockery DW, Pope CA, Xu X, Spengler JD, Ware JH, et al. 1993. An association between air pollution and mortality in six U.S. cities. *N. Engl. J. Med.* 329:1753–59
32. Doll R, Hill AB. 1954. The mortality of doctors in relation to their smoking habits; a preliminary report. *BMJ* 1:1451–55
33. Dominici F, Peng RD, Zeger SL, White RH, Samet JM. 2007. Particulate air pollution and mortality in the United States: Did the risks change from 1987 to 2000? *Am. J. Epidemiol.* 166:880–88
34. Dominici F, Zigler C. 2017. Best practices for gauging evidence of causality in air pollution epidemiology. *Am. J. Epidemiol.* 186:1303–9
35. Dunn C, Kingham S. 1996. Establishing links between air quality and health: searching for the impossible? *Soc. Sci. Med.* 42:831–41
36. Ebenstein A, Fan M, Greenstone M, He G, Zhou M. 2017. New evidence on the impact of sustained exposure to air pollution on life expectancy from China's Huai River Policy. *PNAS* 114:10384–89
37. Farhani S, Ozturk I. 2015. Causal relationship between $CO_2$ emissions, real GDP, energy consumption, financial development, trade openness, and urbanization in Tunisia. *Environ. Sci. Pollut. Res.* 22:15663–76
38. Ferraro PJ, Hanauer MM. 2014. Quantifying causal mechanisms to determine how protected areas affect poverty through changes in ecosystem services and infrastructure. *PNAS* 111:4332–37
39. Fiorito G, Vlaanderen J, Polidoro S, Gulliver J, Galassi C, et al. 2018. Oxidative stress and inflammation mediate the effect of air pollution on cardio- and cerebrovascular disease: a prospective study in nonsmokers. *Environ. Mol. Mutagen.* 59:234–46
40. Flanders WD, Klein M, Darrow LA, Strickland MJ, Sarnat SE, et al. 2011. A method for detection of residual confounding in time-series and other observational studies. *Epidemiology* 22:59–67
41. Frangakis CE, Rubin DB. 2002. Principal stratification in causal inference. *Biometrics* 58:21–29
42. Gamble JF. 1998. PM2.5 and mortality in long-term prospective cohort studies: cause-effect or statistical associations? *Environ. Health Perspect.* 106:535–49
43. Gamble JF, Lewis RJ. 1996. Health and respirable particulate (PM10) air pollution: a causal or statistical association? *Environ. Health Perspect.* 104:838–50
44. Golan R, Kloog I, Almog R, Gesser-Edelsburg A, Negev M, et al. 2018. Environmental exposures and fetal growth: the Haifa pregnancy cohort study. *BMC Public Health* 18:132
45. Goldberg MS. 2007. On the interpretation of epidemiological studies of ambient air pollution. *J. Expo. Sci. Environ. Epidemiol.* 17(Suppl. 2):66–70
46. Goodman JE, Prueitt RL, Sax SN, Bailey LA, Rhomberg LR. 2013. Evaluation of the causal framework used for setting national ambient air quality standards. *Crit. Rev. Toxicol.* 43:829–49
47. Greenwood E. 1945. *Experimental Sociology: A Study in Method*. New York: Octogon Books King's Crown Press
48. Griebel L, Prokosch HU, Köpcke F, Toddenroth D, Christoph J, et al. 2015. A scoping review of cloud computing in healthcare. *BMC Med. Inform. Decis. Mak.* 15:17
49. Habermann M, Gouveia N. 2014. Socioeconomic position and low birth weight among mothers exposed to traffic-related air pollution. *PLOS ONE* 9:e113900
50. Hackstadt AJ, Matsui EC, Williams DL, Diette GB, Breysse PN, et al. 2014. Inference for environmental intervention studies using principal stratification. *Stat. Med.* 33:4919–33
51. Hansen BB, Klopfer SO. 2006. Optimal full matching and related designs via network flows. *J. Comput. Graph. Stat.* 15:609–27
52. Hernberg S, Tolonen M, Nurminen M. 1976. Eight-year follow-up of viscose rayon workers exposed to carbon disulfide. *Scand. J. Work Environ. Health* 2:27–30
53. Hertz-Picciotto I, Schmidt RJ, Krakowiak P. 2018. Understanding environmental contributions to autism: causal concepts and the state of science. *Autism Res.* 11:554–86
54. Hill AB. 1965. The environment and disease: association or causation? *Proc. R. Soc. Med.* 58:295–300

55. Holland PW. 1986. Statistics and causal inference. *J. Am. Stat. Assoc.* 81:945–60
56. Hüls A, Vierkötter A, Sugiri D, Abramson MJ, Ranft U, et al. 2018. The role of air pollution and lung function in cognitive impairment. *Eur. Respir. J.* 51:1701963
57. Imbens GW, Rubin DB. 2015. *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. Cambridge, UK: Cambridge Univ. Press
58. Janes H, Dominici F, Zeger SL. 2007. Trends in air pollution and mortality: an approach to the assessment of unmeasured confounding. *Epidemiology* 18:416–23
59. Jarjour S, Jerrett M, Westerdahl D, de Nazelle A, Hanning C, et al. 2013. Cyclist route choice, traffic-related air pollution, and lung function: a scripted exposure study. *Environ. Health* 12:14
60. Jiang L, Bai L. 2018. Spatio-temporal characteristics of urban air pollutions and their causal relationships: evidence from Beijing and its neighboring cities. *Sci. Rep.* 8:1279
61. Knol AB, de Hartog JJ, Boogaard H, Slottje P, van der Sluijs JP, et al. 2009. Expert elicitation on ultrafine particles: likelihood of health effects and causal pathways. *Part Fibre Toxicol*. 6:19
62. Künzli N, Braun-Fahrländer C, Rapp R, Ackermann-Liebich U. 1997. [Air pollution and health—causal criteria in environmental epidemiology]. *Schweiz. Med. Wochenschr.* 127(33):1334–44 (in German)
63. LaLonde RJ. 1986. Evaluating the econometric evaluations of training programs with experimental data. *Am. Econ. Rev.* 76:604–20
64. Le Tertre A, Quénel P, Eilstein D, Medina S, Prouvost H, et al. 2002. Short-term effects of air pollution on mortality in nine French cities: a quantitative summary. *Arch. Environ. Health* 57:311–19
65. Leal C, Bean K, Thomas F, Chaix B. 2012. Multicollinearity in associations between multiple environmental features and body weight and abdominal fat: using matching techniques to assess whether the associations are separable. *Am. J. Epidemiol.* 175:1152–62
66. Lewis C, Hoggatt KJ, Ritz B. 2011. The impact of different causal models on estimated effects of disinfection by-products on preterm birth. *Environ. Res.* 111:371–76
67. Lu JG, Lee JJ, Gino F, Galinsky AD. 2018. Polluted morality: Air pollution predicts criminal activity and unethical behavior. *Psychol. Sci.* 29:340–55
68. Makar M, Antonelli J, Di Q, Cutler D, Schwartz J, Dominici F. 2017. Estimating the causal effect of low levels of fine particulate matter on hospitalization. *Epidemiology* 28:627–34
69. Marcus AH, Kegler SR. 2001. Confounding in air pollution epidemiology: When does two-stage regression identify the problem? *Environ. Health Perspect.* 109:1193–96
70. Mauderly JL, Kracko D, Brower J, Doyle-Eisele M, McDonald JD, et al. 2014. The National Environmental Respiratory Center (NERC) experiment in multi-pollutant air quality health research: IV. Vascular effects of repeated inhalation exposure to a mixture of five inorganic gases. *Inhal. Toxicol.* 26:691–96
71. Mauderly JL, Seilkop SK. 2014. The National Environmental Respiratory Center (NERC) experiment in multi-pollutant air quality health research: III. Components of diesel and gasoline engine exhausts, hardwood smoke and simulated downwind coal emissions driving non-cancer biological responses in rodents. *Inhal. Toxicol.* 26:668–90
72. Moore KL, Neugebauer R, van der Laan MJ, Tager IB. 2012. Causal inference in epidemiological studies with strong confounding. *Stat. Med.* 31:1380–404
73. Morgan KL, Rubin DB. 2012. Rerandomization to improve covariate balance in experiments. *Ann. Statist.* 40:1263–82
74. Oulhote Y, Bind M-A, Coull B, Patel CJ, Grandjean P. 2017. Combining ensemble learning techniques and g-computation to investigate chemical mixtures in environmental epidemiology studies. bioRxiv 147413. **https://doi.org/10.1101/147413**
75. Owens EO, Patel MM, Kirrane E, Long TC, Brown J, et al. 2017. Framework for assessing causality of air pollution-related health effects for reviews of the National Ambient Air Quality Standards. *Regul. Toxicol. Pharmacol.* 88:332–37
76. Padula AM, Mortimer K, Hubbard A, Lurmann F, Jerrett M, Tager IB. 2012. Exposure to traffic-related air pollution during pregnancy and term low birth weight: estimation of causal associations in a semiparametric model. *Am. J. Epidemiol.* 176:815–24

77. Perret JL, Bowatte G, Lodge CJ, Knibbs LD, Gurrin LC, et al. 2017. The dose-response association between nitrogen dioxide exposure and serum interleukin-6 concentrations. *Int. J. Mol. Sci.* 18:E1015
78. Peyre G, Cuturi M. 2018. Computational optimal transport. arXiv 1803.00567
79. Picciotto S, Eisen EA, Chevrier J. 2014. 0351 g-estimation: Why does it work and what does it offer? *Occup. Environ. Med.* 71:A120–21
80. Pope CA 3rd, Burnett RT. 2007. Confounding in air pollution epidemiology: the broader context. *Epidemiology* 18:424–26
81. Power MC, Adar SD, Yanosky JD, Weuve J. 2016. Exposure to air pollution as a potential contributor to cognitive function, cognitive decline, brain imaging, and dementia: a systematic review of epidemiologic research. *Neurotoxicology* 56:235–53
82. Prandota J, Stolarczyk J. 2002. Autoimmune hepatitis associated with the odour of fish food proteins: a causal relationship or just a mere association? *Allergol. Immunopathol.* 30:331–37
83. Relton CL, Davey Smith G. 2015. Mendelian randomization: applications and limitations in epigenetic studies. *Epigenomics* 7:1239–43
84. Robins J. 1986. A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect. *Math. Model.* 7:1393–512
85. Rosenbaum PR. 2012. Optimal matching of an optimally chosen subset in observational studies. *J. Comput. Graph. Stat.* 21:57–71
86. Rubin DB. 1974. Estimating causal effects of treatments in randomized and nonrandomized studies. *J. Educ. Psychol.* 66:688–701
87. Rubin DB. 2006. William G. Cochran's Contributions to the Design, Analysis, and Evaluation of Observational Studies. In *Matched Sampling for Causal Effects*, pp. 7–29. Cambridge, UK: Cambridge Univ. Press
88. Rubin DB. 2007. The design versus the analysis of observational studies for causal effects: parallels with the design of randomized trials. *Stat. Med.* 26:20–36
89. Saber AT, Jacobsen NR, Jackson P, Poulsen SS, Kyjovska ZO, et al. 2014. Particle-induced pulmonary acute phase response may be the causal link between particle inhalation and cardiovascular disease. *Wiley Interdiscip. Rev. Nanomed. Nanobiotechnol.* 6:517–31
90. Samet JM, Zeger SL, Dominici F, Curriero F, Coursac I, et al. 2000. The National Morbidity, Mortality, and Air Pollution Study. Part II: Morbidity and mortality from air pollution in the United States. *Res. Rep. Health Eff. Inst.* 94:5–70
91. Schwartz J. 1994. Air pollution and hospital admissions for the elderly in Detroit, Michigan. *Am. J. Respir. Crit. Care Med.* 150:648–55
92. Schwartz J. 1994. What are people dying of on high air pollution days? *Environ. Res.* 64:26–35
93. Schwartz J. 1999. Air pollution and hospital admissions for heart disease in eight U.S. counties. *Epidemiology* 10:17–22
94. Schwartz J, Austin E, Bind MA, Zanobetti A, Koutrakis P. 2015. Estimating causal associations of fine particles with daily deaths in Boston. *Am. J. Epidemiol.* 182:644–50
95. Schwartz J, Bind MA, Koutrakis P. 2017. Estimating causal effects of local air pollution on daily deaths: effect of low levels. *Environ. Health Perspect.* 125:23–29
96. Selvin S, Merrill D, Wong L, Sacks ST. 1984. Ecologic regression analysis and the study of the influence of air quality on mortality. *Environ. Health Perspect.* 54:333–40
97. Sheldon TL, Sankaran C. 2017. The impact of Indonesian forest fires on Singaporean pollution and health. *Am. Econ. Rev.* 107:526–29
98. Sinharay R, Gong J, Barratt B, Ohman-Strickland P, Ernst S, et al. 2018. Respiratory and cardiovascular responses to walking down a traffic-polluted road compared with walking in a traffic-free area in participants aged 60 years and older with chronic lung or heart disease and age-matched healthy controls: a randomised, crossover study. *Lancet* 391:339–49
99. Snow J. 1856. On the mode of communication of cholera. *Edinb. Med. J.* 1:668–70
100. Sommer A, Lee M, Bind M-A. 2018. Comparing apples to apples: an environmental criminology analysis of the effects of heat and rain on violent crimes in Boston. *Palgrave Commun.* 4:138

101. Stuart EA. 2010. Matching methods for causal inference: a review and a look forward. *Stat. Sci.* 25:1–21
102. Thistlethwaite DL, Campbell DT. 1960. Regression-discontinuity analysis: an alternative to the ex post facto experiment. *J. Educ. Psychol.* 2:119–28
103. Tolonen M, Hernberg S, Nurminen M, Tiitola K. 1975. A follow-up study of coronary heart disease in viscose rayon workers exposed to carbon disulphide. *Br. J. Ind. Med.* 32:1–10
104. Tommasi S, Zheng A, Yoon J-I, Li AX, Wu X, Besaratinia A. 2012. Whole DNA methylome profiling in mice exposed to secondhand smoke. *Epigenetics* 7:1302–14
105. Volk HE, Lurmann F, Penfold B, Hertz-Picciotto I, McConnell R. 2013. Traffic-related air pollution, particulate matter, and autism. *JAMA Psychiatry* 70:71–77
106. Wang Y, Kloog I, Coull BA, Kosheleva A, Zanobetti A, Schwartz JD. 2016. Estimating causal effects of long-term PM2.5 exposure on mortality in New Jersey. *Environ. Health Perspect.* 124:1182–88
107. Weisskopf MG, Kioumourtzoglou MA, Roberts AL. 2015. Air pollution and autism spectrum disorders: causal or confounded? *Curr. Environ. Health Rep.* 2:430–39
108. Wong C-M, Atkinson RW, Anderson HR, Hedley AJ, Ma S, et al. 2002. A tale of two cities: effects of air pollution on hospital admissions in Hong Kong and London compared. *Environ. Health Perspect.* 110:67–77
109. Woodruff PG, Ellwanger A, Solon M, Cambier CJ, Pinkerton KE, Koth LL. 2009. Alveolar macrophage recruitment and activation by chronic second hand smoke exposure in mice. *COPD* 6:86–94
110. Yang C, Zhao W, Deng K, Zhou V, Zhou X, Hou Y. 2017. The association between air pollutants and autism spectrum disorders. *Environ. Sci. Pollut. Res. Int.* 24:15949–58
111. Yang S, Tan Y, Mei H, Wang F, Li N, et al. 2018. Ambient air pollution the risk of stillbirth: a prospective birth cohort study in Wuhan, China. *Int. J. Hyg. Environ. Health* 221:502–9
112. Zhong J, Trevisi L, Urch B, Lin X, Speck M, et al. 2017. B-vitamin supplementation mitigates effects of fine particles on cardiac autonomic dysfunction and inflammation: a pilot human intervention trial. *Sci. Rep.* 7:45322
113. Zigler CM, Choirat C, Dominici F. 2018. Impact of National Ambient Air Quality Standards nonattainment designations on particulate pollution and health. *Epidemiology* 29:165–74
114. Zigler CM, Dominici F. 2014. Point: clarifying policy evidence with potential-outcomes thinking—beyond exposure-response estimation in air pollution epidemiology. *Am. J. Epidemiol.* 180:1133–40
115. Zigler CM, Dominici F, Wang Y. 2012. Estimating causal effects of air quality regulations using principal stratification for spatially correlated multivariate intermediate outcomes. *Biostatistics* 13:289–302
116. Zigler CM, Kim C, Choirat C, Hansen JB, Wang Y, et al. 2016. *Causal inference methods for estimating long-term health effects of air quality regulations.* Res. Rep. 187, Health Eff. Inst. **https://www.healtheffects.org/system/files/ZiglerRR187.pdf**
117. Zubizarreta JR, Kilcioglu C, Vielma JP. 2018. designmatch: matched samples that are balanced and representative by design. *R Package Version 0.3.1.* **https://cran.r-project.org/web/packages/designmatch/index.html**