In Pursuit of Evidence in Air Pollution Epidemiology: The Role of Causally Driven Data Science

Marco Carone[a], Francesca Dominici[b], and Lianne Sheppard[a,c]

The impact of air pollution on health gained widespread attention in 1952 when pollution from coal burning in London, with smoke so thick that drivers needed headlights at all times of day, was linked to a dramatic rise in the number of deaths.[1,2] Over the past decades, evidence of adverse population health effects has accumulated from many thousands of epidemiologic studies, suggesting this is a critically important public health problem. While at the individual level the relative effects of exposure tend to be small, and potentially large confounding biases are difficult to rule out, air pollution is ubiquitous and the exposed population is enormous. Thus, these adverse effects have important policy implications and intense scrutiny of these epidemiologic findings is inevitable.[3,4] Recently, the emergence onto the statistics landscape of modern causal inference methods, as well as machine learning and other novel data science techniques, has generated excitement and a sense of promise that estimation of the causal effects of air pollution exposures on health outcomes could be more definitively described.[5–7] (By causal effects we mean the effects that would be seen under experimental changes of exposures. By causal inference we mean the process of inferring causal effects from data.) We share the excitement of others that the discipline of causal inference has the potential to advance air pollution policy and allow the integration of modern statistical tools into air pollution epidemiology, but we also caution against unrealistic expectations by highlighting important difficulties ahead. Our goal is to provide a glimpse of the opportunities afforded by the use of causal inference and data science methods, raise awareness about some of the outstanding challenges, and inspire others to join on the efforts to overcome them. In our note, we first discuss broad concepts that are relevant across applications, and then focus issues germane to air pollution epidemiology.

THERE IS A NEED FOR DELIBERATE CAUSAL INFERENCE IN POLICY-RELEVANT RESEARCH

When the goal of a scientific study is to inform policy, describing associations between exposures and outcomes generally does not suffice: an assessment of causal effects is needed.[8] Unless a particular study design very clearly allows such an assessment, as in an appropriately designed and conducted randomized trial, epidemiologists and biostatisticians typically word their findings carefully to avoid implying causation.[9] Nevertheless, since the premise of these analyses is to bring to light potentially causal relationships, despite cautious wording, investigators and end users both may implicitly infer causation. For this reason, we believe it is important for analysts to employ methodology that allows causal inference under the most realistic set of conditions possible, and to ensure clear communication of these conditions to facilitate the interpretation of the results.

CONVENTIONAL MODEL-BASED APPROACHES MAY BE INADEQUATE FOR INFERRING CAUSAL EFFECTS

In many disciplines, conventional regression models (e.g., the very many varieties of the linear model) are considered a cornerstone of traditional statistical practice for inference about health effects. Indeed,

From the [a]Department of Biostatistics, University of Washington, Seattle, WA; [b]Department of Biostatistics, Harvard T. H. Chan School of Public Health, Harvard University, Boston, MA; and [c]Department of Environmental and Occupational Health Sciences, University of Washington, Seattle, WA.

L.S. and M.C. were supported by NIH grant R01 ES026187. F.D. was supported by the Health Effects Institute (HEI), an organization jointly funded by the United States Environmental Protection Agency (EPA) (Assistance Award No.CR-83467701) and certain motor vehicle and engine manufacturers. The contents of this article do not necessarily reflect the views of HEI, or its sponsors, nor do they necessarily reflect the views and policies of the EPA or motor vehicle and engine manufacturers. This commentary was sponsored by the International Society of Environmental Epidemiology (ISEE). The contents are the sole responsibility of the author(s) and do not necessarily reflect the official views of the ISEE.

The authors report no conflicts of interest.

Correspondence: Lianne Sheppard, Box 357232, Department of Biostatistics, University of Washington, Seattle, WA 98195-7232. E-mail: sheppard@uw.edu.

ISSN: 1044-3983/20/3101-0001
DOI: 10.1097/EDE.0000000000001090

the great majority of air pollution epidemiology studies estimate health effects as relative risks or hazard ratios obtained using regression models. In the conventional model-based approach, it is customary to begin by selecting a particular regression model, often on the basis of the type of outcome under study (e.g., logistic regression for binary outcomes and proportional hazards model for survival outcomes). How notions of causality may be ascribed is then considered as a second step, often restricted to consideration of which potential confounders ought to be included in the model. In this conventional approach, investigators often focus on parameters most conveniently reported on the basis of the model chosen (e.g., odds ratio in logistic regression and hazard ratio in proportional hazards model).

Unfortunately, this common and traditional use of conventional regression models for causal inference has at least two important drawbacks. First, whether the estimation procedure used yields valid inferences on a scientifically meaningful estimand (i.e., the quantity being estimated) – let alone one with a causal interpretation – generally hinges on the assumption that the specified regression model accurately reflects the relationship modeled.[10–12] And that the true underlying mechanism follows the simple form of conventional regression models is often a strong assumption without much prior empirical support. Additionally, well-intentioned efforts to reduce model misspecification through an iteration of model revisions guided by diagnostics (e.g., residual plots) can compromise the interpretation and validity of inferences from these models.[13–15] Accounting for informal model selection in model-based analyses is difficult and remains an unsettled topic of methodologic research.[16–20] Second, even in the ideal scenario in which all relevant confounders are accounted for and the regression model postulated a priori holds true, model-based regression coefficients corresponding to the exposure of interest may still not refer to the causal contrast desired to address the scientific question at hand. This may occur, for example, because the regression coefficients quantify a causal effect on a different scale than desired (e.g., odds ratio vs. relative risk) or do not provide the population-level summary desired (e.g., conditional vs. marginal interpretation). The limitations of conventional model-based causal inference are exacerbated when modeling is performed on scales less amenable to causal comparisons (e.g., the hazard ratio), or when the exposure of interest occurs over a longer period of time.[21–25]

HOW DO MODERN CAUSAL INFERENCE METHODS DIFFER?

Modern causal inference methods help circumvent the limitations of conventional model-based approaches. First, as we discuss below, these methods facilitate the use of more flexible learning approaches. As such, they eliminate the overreliance of conventional model-based approaches on strong model assumptions, thereby increasing the reliability of the resulting scientific findings. Second, in modern causal inference, the causal estimand is rigorously defined before the inferential method is determined. Because it is deliberately chosen rather than inherited from the model choice that the outcome data type might suggest, the estimand is more likely to be directly relevant to address the scientific question. These points are discussed further below.

The choice of causal estimand plays a fundamental role as the starting point of all statistical considerations in modern causal inference methods. Counterfactual (or potential) outcomes, which are used to refer to hypothetical versions of the outcome under different exposure profiles of interest, play a key role in this first step. Causal estimands are generally expressed as summaries of the distribution of counterfactuals.[26] For example, in a simple binary exposure setting, the average causal effect is the mean value of the difference between the counterfactual outcomes corresponding to each of the two exposure profiles (e.g., exposed at high vs. low levels). These causal estimands cannot be directly estimated because only a single counterfactual can be observed for each study participant – for example, for an individual having been exposed at a high exposure level, only the counterfactual outcome corresponding to high exposure is observed. To identify causal estimands from the observed data, there must exist an appropriate "bridge" (i.e., mapping) from the counterfactual world, where causal estimands are defined, to the observed world, in which data are collected. This bridge must allow the causal estimand to be expressed as a statistical estimand, that is, as a summary of the distribution of the observed data (Figure 1). The existence of a bridge hinges on meeting the causal conditions, many of which are untestable. For the average causal effect, typical causal conditions used for identification include the following: consistency (i.e., the intervention defining exposures is unambiguous); the stable unit treatment value assumption (i.e., no interference); positivity (i.e., each participant could have been observed in the counterfactual exposure group); and (sequential) ignorability (i.e., exchangeability or no unmeasured confounding).[26] When such a bridge exists, the causal estimand is said to be identified.[26] For example, if the average causal effect (say of a high vs. low binary exposure) is the causal estimand of interest, the destination of one such bridge is the G-computation formula. The latter is given by the population average of the difference in strata-specific mean outcomes among individuals with high vs. low exposure, where strata correspond to subpopulations defined by confounder levels.[26] We invite readers to consult the textbook of Hernán and Robins for an in-depth study of foundational ideas in causal inference.[26] For a comprehensive discussion of causal inference in air pollution epidemiology, we point readers to the recent review by Bind.[7]

Once identification of the causal estimand is established, the resulting statistical parameter can be estimated using a variety of strategies. At this point, the problem is purely statistical: the fact that the statistical parameter corresponds to the causal estimand under certain causal conditions is irrelevant to how estimation and inference should then proceed. Simplifying statistical conditions, such as parametric forms (e.g., linear mean regression model), can be imposed on the observed data distribution to facilitate estimation and inference, but they are not strictly needed. This is where flexible learning techniques, such as machine learning, can be deployed to perform statistical inference less prone to bias due to model misspecification. Prior knowledge, including characteristics of the study design, informs all the conditions and can justify imposing more stringent statistical conditions that can be leveraged to generate more precise inference. Of course, whether or not a causal interpretation may be ascribed to the inferences drawn depends on the validity of the causal conditions.

In summary, as we highlight in the figure, a desirable feature of the causal inference framework is that it separates conditions as being either causal or statistical, with the latter generally being imposed out of convenience rather than necessity. In contrast, in the conventional model-based approach, causal and statistical conditions are often entangled. As such, without additional work, it is often difficult to determine whether identification is in fact largely driven by strong statistical modeling conditions (e.g., assumption of linearity of the mean outcome beyond the range of observed exposures[27]) – this reduces transparency and rigor in the conduct of causal inference. Explicitly disentangling causal and statistical conditions is also important because very different strategies exist to relax each set of conditions.

CAUSALLY GUIDED DATA SCIENCE HAS THE POTENTIAL TO ADVANCE POLICY-RELEVANT DECISION-MAKING

Because causal estimands, such as the average causal effect, have clearer scientific interpretability than statistical estimands based

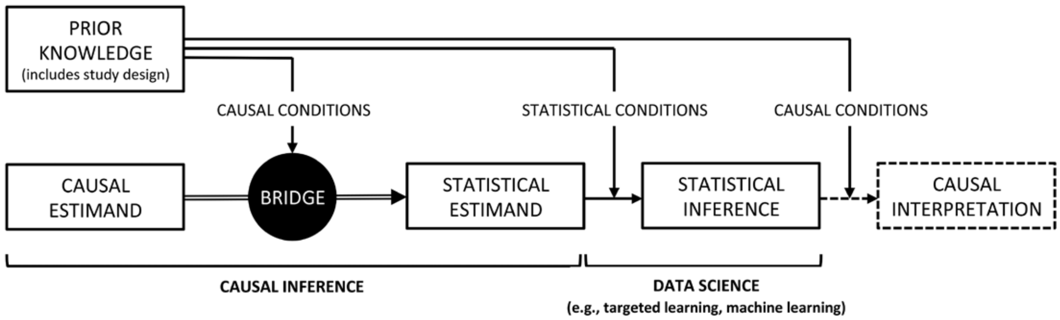

FIGURE 1. Framework for causally guided data science.

on (plausibly misspecified) conventional models (e.g., coefficients in regression models), they are better suited to inform air pollution policy-making.[28] These causal estimands can quantify the health consequences of hypothetical (possibly regulatory-driven) changes in exposure, and can predict the effects of future interventions or policies.[5] Modern causal inference methods therefore widen the scope and increase the granularity of policy-relevant scientific questions that can be addressed from the available data.[5] Additionally, reliable inference in air pollution epidemiology is needed, particularly in view of the wide-ranging implications of air pollution policies. Results from analyses that build upon unrealistic conditions may be misleading.

Causal conditions allow identification of the causal parameter without generally implying constraints on the observed data distribution. As discussed above, the use of conventional statistical models, such as conventional parametric models, often impose unnecessary conditions for the validity of inferences, regardless of whether the causal conditions are met. This is where we believe developments in data science can have the greatest impact. Statisticians and other data scientists have developed many flexible learning algorithms to reliably and robustly uncover structure from data. Their use, however, does not imply that more conventional learning methods (e.g., linear regression) must be discarded. Certain ensemble learning methods (e.g., the Super Learner, an example of model stacking) are able to combine algorithms in any user-specified collection of learning strategies in an automatic, data-driven, and optimal fashion.[29,30] This eliminates the need to bet on any particular learning tool and also largely negates the relevance of model diagnostics. To perform optimal inference for the statistical parameter of interest that identifies the causal estimand, the output of any flexible learning method must generally be corrected to achieve an optimal bias-variance trade-off for the desired estimand. There are several possible strategies for doing so, some of which have been used for decades (e.g., estimating equations methodology) and others are more recent innovations (e.g., targeted minimum loss-based estimation).[31–33] What emerges is thus a two-step process, wherein relevant features of the observed data distribution are estimated flexibly and then corrected to ensure valid inference for the actual target of inference. The study of the tools and relevant issues to consider in this process has been the focus of targeted learning, a rapidly expanding area of methodologic research.[32,34–39] These tools allow practitioners to employ flexible learning strategies, thus facilitating robust analyses based on as few statistical conditions as possible. They also enable reliable uncertainty assessment, including confidence intervals and *P*-values, a critical ingredient in the interpretation of results for the sake of policy-making.

When causal conditions fail to hold (e.g., there is interference or unmeasured confounding), a bridge may not exist, and full identification of the causal estimand is then not possible. In some cases, partial identification, wherein bounds on the true causal estimand are identified, holds.[40] Full identification can be recovered by modifying the definition of the exposure or the target population. An alternative strategy is to find an instrument – a cause of exposure that otherwise has no bearing on the outcome – to help circumvent failure to account for all important confounders, including the ones that are unmeasured.[41,42] Examples include quasi-experiments such as the coal ban in Dublin and the traffic plan implemented during the Beijing Olympics.[43,44] Regardless, even when causal conditions do not strictly hold, causally-motivated estimands still have a transparent, model-agnostic interpretation that is often more sensible than that of conventional model-based estimands. We believe this enables more informed policy-relevant decision making about air pollution.

CAUSAL INFERENCE IN AIR POLLUTION EPIDEMIOLOGY IS INHERENTLY CHALLENGING

Investigating the causal effect of exposure to air pollution on health outcomes is an intrinsically challenging endeavor.[45,46] We highlight only a few of these challenges below.

First, obvious ethical considerations often preclude the conduct of randomized trials, a design that allows easy derivation of causal inferences. Observational studies face confounding biases that may be difficult to fully account for. Furthermore, these biases can be in either direction and can have unknown magnitude. This challenge is magnified because for individuals the relative increase in risk from air pollution exposures is typically small.[47] Of course, this is not to say that these adverse effects should be overlooked – ubiquitous air pollution exposures can still affect individuals, and, more importantly, accumulate over large populations, resulting in a substantial burden for communities and health systems.[48] Uncontrolled confounding results in violation of ignorability, a key causal condition, and is difficult to avoid air pollution epidemiology.

Second, when the health effects arise from long-term exposures, defining causal effects of interest can be challenging because of the many possible definitions of counterfactual outcomes. Defining the counterfactual outcomes requires selection of the relevant interventions on exposures. What exposure duration should be considered? Should daily exposure be fixed at a particular level across time, or allowed to vary within a window? Being more or less strict in defining patterns of exposure leads to a trade-off between the interpretability of causal effects and the feasibility of the statistical problem. Oversimplification of the complex exposure process – for example, by ignoring the time-varying nature of exposures – can potentially invalidate causal inferences, as some authors have recently argued.[49] Marginal structural models, a set of tools for performing parsimonious causal regression, and causal estimands based on stochastic interventions seem particularly promising to tackle these challenges.[25,50–55]

Third, even when using causal inference methods, disentangling the causal effects of exposure to a particular pollutant may be difficult if it tends to co-occur with other pollutants.[56–63] This may be especially true if these pollutants emanate from the same source. Since the co-occurring exposures possibly confound the relationship between the pollutant and health outcome of interest, and must therefore be accounted for, positivity violations are likely. Learning accurately how the outcome differs based jointly on the various exposures and confounders is then difficult, since the health effects are at best only weakly identified.

Fourth, accurately measuring exposure to air pollutants is itself a challenging task.[47,64] While technical advances have made it possible to outfit individuals with personal sensors capable of measuring exposure histories, implementing this in large-scale studies is prohibitively costly. Instead, by combining data obtained from regulatory monitoring stations, geographic information systems, and satellite imaging, researchers have developed historical exposure models used to predict pollutant levels in space and time, particularly at unmonitored locations.[65–69] However sophisticated these models may be, their outputs remain estimates. Accounting for their inherent (spatially varying) uncertainty and biases largely remains an unresolved problem in air pollution epidemiology, although recent advances have been made.[70–75] Additionally, even ignoring the imperfect mapping of historical pollutant levels, individual location tracking would be needed to obtain accurate exposure histories. Currently, most cohort studies use outdoor pollutant levels at domicile location as a proxy to reconstruct exposure history, ignoring the fact that most exposures may occur in an entirely different locale. It is important to understand the repercussions of these issues on causal inference and to identify possible remedies.[75]

Fifth, causal methods are emerging and their development remains the focus of active research. Considerable advances in methodological research and software tools are likely needed before many practitioners will be ready to adopt them as standard practice. Challenges include the following: determining appropriate counterfactuals; spelling out relevant causal estimands; determining identification formulas (i.e., relating these causal estimands to summaries of the data-generating distribution – the bridge); and devising inferential procedures that leverage flexible, data science tools (e.g., machine learning). These tasks require substantial training in relatively recent technical areas, such as causal inference, targeted learning, and machine learning, which many academic programs do not yet cover in depth. The unfamiliarity of these modern tools to many scientists may also be a perceived barrier to publication, possibly driving researchers to persist in the use of more conventional methods.

RESPONSIBLE POLICY-MAKING SHOULD BE INFORMED BUT NOT PARALYZED BY CAUSAL INFERENCE

We view the greater adoption of cutting-edge data science tools and causal inference principles into mainstream air pollution epidemiology as an important step forward.[5,8,76–84] As Goldman and Dominici state, "well-validated methods for causal inference can play a useful role: this is because they include a more transparent disclosure of all the assumptions that are needed to properly adjust for confounding compared with regression modeling and therefore can infer causality in analyses of observational data."[85] Nevertheless, we cannot expect the emergence of a methodologic silver bullet since many of the challenges of drawing valid causal inferences about air pollution health effects stem from inherent features of the observational nature of the available data. In other words, no statistical approach is likely to overcome these challenges entirely. Rather, cutting-edge methods must be combined with the use of innovative study designs or complementary data sources tailored to the particular difficulties encountered in air pollution epidemiology.

Some scientists have argued that, because epidemiologic studies provide measures of association and may not accurately predict the benefits of reducing air pollution, they should not be used, thereby dismissing a large body of evidence painstakingly gathered over decades.[86,87] Yet, the Clean Air Act mandates that, despite the presence of uncertainty, air quality regulations can be set providing for an adequate margin of safety.[88,89] In our view, causal inference methods should not be used as another opportunity to weaponize science against itself. Policymakers cannot wait for the data, study designs, and analytic tools that will ensure unarguable causal inferences: stalling until perfect evidence arises is irresponsible and does not protect public health.

ACKNOWLEDGMENTS

This commentary was originally presented to kick off the discussion at the "Causal Modeling in Air Pollution Research and Policy" preconference workshop at the 2018 HEI Annual Conference. The authors wish to thank Katherine Walker, Sverre Vedal, Rachel Shaffer, Patrick Heagerty, and Mark van der Laan for helpful comments on earlier drafts of this work, and three anonymous referees for their constructive feedback.

ABOUT THE AUTHORS

MARCO CARONE is an assistant professor of biostatistics at the University of Washington School of Public Health, and an Affiliate Investigator in the Vaccine and Infectious Disease Division at the Fred Hutchinson Cancer Research Center. His research lies at the intersection of nonparametric and semiparametric statistics, causal inference, survival analysis, and statistical epidemiology. He focuses on developing efficient and robust methods for performing statistical inference with machine learning tools. He is the Norman Breslow Endowed Faculty Fellow.

FRANCESCA DOMINICI is a professor of biostatistics at the Harvard T.H. Chan School of Public Health and Co-Director of the Data Science Initiative at Harvard University. Her research focuses on developing and advancing methods for the analysis of large, heterogeneous data sets to identify and understand the health impacts of environmental threats and inform policy. She is a member of the National Academy of Medicine.

LIANNE SHEPPARD is a professor of biostatistics, and environmental and occupational health sciences at the University of Washington School of Public Health. Her research interests focus on statistical methods for understanding the health effects of environmental and occupational exposures; they include study design, measurement error, exposure modeling and estimation, and estimation of environmental exposure effects with application to a wide range of health outcomes. She is a fellow of the American Statistical Association and served on the chartered Clean Air Scientific Advisory Committee.

REFERENCES

1. Logan WPD. Mortality in the London fog incident, 1952. *Lancet*. 1953;264:336–338.
2. Bell ML, Davis DL. Reassessment of the lethal London fog of 1952: novel indicators of acute and chronic consequences of acute exposure to air pollution. *Environ Health Perspect*. 2001;109:389–394.

3. Taubes G. Epidemiology faces its limits. *Science*. 1995;269:164–169.
4. Ayres JG. Health effects of gaseous air pollutants. *Issues in Environmental Science and Technology*. 1998;10:1–20.
5. Zigler CM, Dominici F. Point: clarifying policy evidence with potential-outcomes thinking-beyond exposure-response estimation in air pollution epidemiology. *Am J Epidemiol*. 2014;180:1133–1140.
6. Hubbell B, Greenbaum D. Counterpoint: moving from potential-outcomes thinking to doing–changing research planning to enable successful health outcomes research. *Am J Epidemiol*. 2014;180:1141–1144.
7. Bind M-A. Causal modeling in environmental health. *Annu Rev Public Health*. 2019;40:23–43.
8. Dominici F, Zigler C. Best practices for gauging evidence of causality in air pollution epidemiology. *Am J Epidemiol*. 2017;186:1303–1309.
9. Hernán MA. The C-word: scientific euphemisms do not improve causal inference from observational data. *Am J Public Health*. 2018;108:616–619.
10. Struthers CA, Kalbfleisch JD. Misspecified proportional hazard models. *Biometrika*. 1986;73:363–369.
11. Lin DY, Wei LJ. The robuest inference for the cox proportional hazards model. *J Am Stat Assoc*. 1989;84:1074–1078.
12. Dominici F, Greenstone M, Sunstein CR. Science and regulation. particulate matter matters. *Science*. 2014;344:257–259.
13. Potscher BM. Effects of model selection inference. *Econometric Theory*. 1991;7:163–185.
14. Chatfield C. Model uncertainty, data mining and statistical inference. *J R Stat Soc Ser A*. 1995;158:419–466.
15. Leeb H, Potscher BM. Model selection and inference: facts and fiction. *Econometric Theory*. 2005;21:21–59.
16. Berk R, Brown L, Zhao LD. Statistical inference after model selection. *J Quant Criminol*. 2010;26:217–236.
17. Berk R, Brown L, Buja A, Zhang K, Zhao LD. Valid post-selection inference. *Ann Stat*. 2013;41:802–837.
18. Taylor J, Tibshirani RJ. Statistical learning and selective inference. *Proc Natl Acad Sci USA*. 2015;112:7629–7634.
19. Tian XY, Taylor J. Asymptotics of selective inference. *Scand Stat Theory Appl*. 2017;44:480–499.
20. Chernozhukov V, Hansen C, Spindler M. Valid post-selection and post-regularization inference: an elementary, general approach. *Annu Rev Econom*. 2015;7:649–688.
21. Hernán MA. The hazards of hazard ratios. *Epidemiology*. 2010;21:13–15.
22. Aalen OO, Cook RJ, Røysland K. Does Cox analysis of a randomized survival study yield a causal treatment effect? *Lifetime Data Anal*. 2015;21:579–593.
23. Alexander BM, Schoenfeld JD, Trippa L. Hazards of hazard ratios - deviations from model assumptions in immunotherapy. *N Engl J Med*. 2018;378:1158–1159.
24. Martinussen T, Vansteelandt S, Andersen PK. Subtleties in the interpretation of hazard ratios. *arXiv*. XXX2018:1810.09192.
25. Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology*. 2000;11:550–560.
26. Hernan MA, Robins JM. *Causal Inference*. Boca Raton, Fla.: Chapman & Hall/CRC; 2019, forthcoming.
27. Nethery RC, Mealli F, Dominici F. Estimating population average causal effects in the presence of non-overlap: a bayesian approach. *Ann Appl Stat*. 2019;13:1242–1267.
28. Glass TA, Goodman SN, Hernán MA, Samet JM. Causal inference in public health. *Annu Rev Public Health*. 2013;34:61–75.
29. van der Laan MJ, Polley EC, Hubbard AE. Super learner. *Stat Appl Genet Mol Biol*. 2007;6:Article25.
30. Polley EC, Rose S, Van der Laan MJ. *Super Learning. Targeted Learning: Causal Inference for Observational and Experimental Data*. New York, N.Y.: Springer; 2011:43–66.
31. Pfanzagl J. Lecture notes in statistics. *Contributions to a General Asymptotic Statistical Theory*. New York, Heidelberg, Berlin: Springer-Verlag 1982:13.
32. Van der Laan MJ, Rose S. *Targeted Learning: Causal Inference for Observational and Experimental Data*. New York: Springer Science & Business Media; 2011.
33. Van der Laan MJ, Laan M, Robins JM. *Unified methods for censored longitudinal data and causality*. New York: Springer Science & Business Media; 2003.
34. van der Laan MJ. Targeted maximum likelihood based causal inference: part I. *Int J Biostat*. 2010;6:Article 2.
35. van der Laan MJ. Targeted maximum likelihood based causal inference: part II. *Int J Biostat*. 2010;6:Article 3.
36. Van der Laan MJ, Rose S. *Targeted Learning in Data Science: Causal Inference for Complex Longitudinal Studies*. Springer; 2018.
37. van der Laan MJ, Starmans RJCM. Entering the era of data science: targeted learning and the integration of statistics and computational data analysis. *Advances in Statistics*. 2014;2014;19.
38. Benkeser D, Carone M, Laan MJV, Gilbert PB. Doubly robust nonparametric inference on the average treatment effect. *Biometrika*. 2017;104:863–880.
39. Carone M, Diaz I, van der Laan MJ. Higher-order targeted loss-based estimation. *Targeted Learning in Data Science: Causal Inference for Complex Longitudinal Studies*. Springer international publishing AG; 2018:483–510.
40. Richardson A, Hudgens MG, Gilbert PB, Fine JP. Nonparametric bounds and sensitivity analysis of treatment effects. *Stat Sci*. 2014;29:596–618.
41. Angrist JD, Imbens GW, Rubin DB. Identification of causal effects using instrumental variables. *J Am Stat Assoc*. 1996;91:444–455.
42. Baiocchi M, Cheng J, Small DS. Instrumental variable methods for causal inference. *Stat Med*. 2014;33:2297–2340.
43. Clancy L, Goodman P, Sinclair H, Dockery DW. Effect of air-pollution control on death rates in Dublin, Ireland: an intervention study. *Lancet*. 2002;360:1210–1214.
44. 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;307:2068–2078.
45. Pearce N, Vandenbroucke JP, Lawlor DA. Causal inference in environmental epidemiology: old and new approaches. *Epidemiology*. 2019;30:311–316.
46. Flanders WD, Garber MD. Is the smog lifting?: causal inference in environmental epidemiology. *Epidemiology*. 2019;30:317–320.
47. Sheppard L, Burnett RT, Szpiro AA, et al. Confounding and exposure measurement error in air pollution epidemiology. *Air Qual Atmos Health*. 2012;5:203–216.
48. Fann N, Kim SY, Olives C, Sheppard L. Estimated changes in life expectancy and adult mortality resulting from declining PM2.5 exposures in the contiguous United States: 1980–2010. *Environ Health Perspect*. 2017;125:097003.
49. Etievant L, Viallon V. Causal inference under over-simplified longitudinal causal models. *arXiv*. 2018:1810.01294.
50. Neugebauer R, van der Laan MJ, Joffe MM, Tager IB. Causal inference in longitudinal studies with history-restricted marginal structural models. *Electron J Stat*. 2007;1:119–154.
51. Neugebauer R, van der Laan M. Nonparametric causal effects based on marginal structural models. *J Stat Plan Inference*. 2007;137:419–434.
52. Petersen M, Schwab J, Gruber S, Blaser N, Schomaker M, van der Laan M. Targeted maximum likelihood estimation for dynamic and static longitudinal marginal structural working models. *J Causal Inference*. 2014;2:147–185.
53. Haneuse S, Rotnitzky A. Estimation of the effect of interventions that modify the received treatment. *Stat Med*. 2013;32:5260–5277.
54. Díaz I, van der Laan MJ. Assessing the causal effect of policies: an example using stochastic interventions. *Int J Biostat*. 2013;9:161–174.
55. Kennedy EH. Nonparametric causal effects based on incremental propensity score interventions. *J Am Stat Assoc*. 2019;114:645–656.
56. Jandarov RA, Sheppard LA, Sampson PD, Szpiro AA. A novel principal component analysis for spatially misaligned multivariate air pollution data. *J R Stat Soc Ser C Appl Stat*. 2017;66:3–28.
57. Bobb JF, Valeri L, Claus Henn B, et al. Bayesian kernel machine regression for estimating the health effects of multi-pollutant mixtures. *Biostatistics*. 2015;16:493–508.

58. Liu SH, Bobb JF, Claus Henn B, et al. Bayesian varying coefficient kernel machine regression to assess neurodevelopmental trajectories associated with exposure to complex mixtures. *Stat Med*. 2018;37:4680–4694.
59. Liu SH, Bobb JF, Lee KH, et al. Lagged kernel machine regression for identifying time windows of susceptibility to exposures of complex mixtures. *Biostatistics*. 2018;19:325–341.
60. Kim SY, Sheppard L, Kaufman JD, et al. Individual-level concentrations of fine particulate matter chemical components and subclinical atherosclerosis: a cross-sectional analysis based on 2 advanced exposure prediction models in the multi-ethnic study of atherosclerosis. *Am J Epidemiol*. 2014;180:718–728.
61. Dominici F, Peng RD, Barr CD, Bell ML. Protecting human health from air pollution: shifting from a single-pollutant to a multipollutant approach. *Epidemiology*. 2010;21:187–194.
62. Vedal S, Kaufman JD. What does multi-pollutant air pollution research mean? *Am J Respir Crit Care Med*. 2011;183:4–6.
63. Vedal S, Campen MJ, McDonald JD, et al. National Particle Component Toxicity (NPACT) initiative report on cardiovascular effects. *Res Rep Health Eff Inst*. 2013;178:5–8.
64. Zeger SL, Thomas D, Dominici F, et al. Exposure measurement error in time-series studies of air pollution: concepts and consequences. *Environ Health Perspect*. 2000;108:419–426.
65. Lindström J, Szpiro AA, Sampson PD, et al. A flexible spatio-temporal model for air pollution with spatial and spatio-temporal covariates. *Environ Ecol Stat*. 2014;21:411–433.
66. Yanosky JD, Paciorek CJ, Suh HH. Predicting chronic fine and coarse particulate exposures using spatiotemporal models for the Northeastern and Midwestern United States. *Environ Health Perspect*. 2009;117:522–529.
67. Di Q, Kloog I, Koutrakis P, Lyapustin A, Wang Y, Schwartz J. Assessing PM2.5 Exposures with high spatiotemporal resolution across the continental United States. *Environ Sci Technol*. 2016;50:4712–4721.
68. Kim SY, Olives C, Sheppard L, et al. Historical prediction modeling approach for estimating long-term concentrations of PM2.5 in cohort studies before the 1999 implementation of widespread monitoring. *Environ Health Perspect*. 2017;125:38–46.
69. Keller JP, Olives C, Kim SY, et al. A unified spatiotemporal modeling approach for predicting concentrations of multiple air pollutants in the multi-ethnic study of atherosclerosis and air pollution. *Environ Health Perspect*. 2015;123:301–309.
70. Szpiro AA, Paciorek CJ, Sheppard L. Does more accurate exposure prediction necessarily improve health effect estimates? *Epidemiology*. 2011;22:680–685.
71. Szpiro AA, Sheppard L, Lumley T. Efficient measurement error correction with spatially misaligned data. *Biostatistics*. 2011;12:610–623.
72. Szpiro AA, Paciorek CJ. Measurement error in two-stage analyses, with application to air pollution epidemiology. *Environmetrics*. 2013;24:501–517.
73. Bergen S, Szpiro AA. Mitigating the impact of measurement error when using penalized regression to model exposure in two-stage air pollution epidemiology studies. *Environ Ecol Stat*. 2015;22:601–631.
74. Bergen S, Sheppard L, Kaufman JD, Szpiro AA. Multipollutant measurement error in air pollution epidemiology studies arising from predicting exposures with penalized regression splines. *J R Stat Soc Ser C Appl Stat*. 2016;65:731–753.
75. Wu X, Braun D, Kioumourtzoglou M-A, Choirat C, Di Q, Dominici F. Causal inference in the context of an error prone exposure: air pollution and mortality. *Ann Appl Stat*. 2019;13:520–547.
76. Zigler CM, Dominici F, Wang Y. Estimating causal effects of air quality regulations using principal stratification for spatially correlated multivariate intermediate outcomes. *Biostatistics*. 2012;13:289–302.
77. Zigler CM, Dominici F. Uncertainty in propensity score estimation: bayesian methods for variable selection and model averaged causal effects. *J Am Stat Assoc*. 2014;109:95–107.
78. Zigler CM, Choirat C, Dominici F. Impact of national ambient air quality standards nonattainment designations on particulate pollution and health. *Epidemiology*. 2018;29:165–174.
79. Moore K, Neugebauer R, Lurmann F, et al. Ambient ozone concentrations cause increased hospitalizations for asthma in children: an 18-year study in Southern California. *Environ Health Perspect*. 2008;116:1063–1070.
80. Moore K, Neugebauer R, Lurmann F, et al. Ambient ozone concentrations and cardiac mortality in Southern California 1983-2000: application of a new marginal structural model approach. *Am J Epidemiol*. 2010;171:1233–1243.
81. Schwartz J, Austin E, Bind MA, Zanobetti A, Koutrakis P. Estimating causal associations of fine particles with daily deaths in Boston. *Am J Epidemiol*. 2015;182:644–650.
82. Schwartz J, Bind MA, Koutrakis P. Estimating causal effects of local air pollution on daily deaths: effect of low levels. *Environ Health Perspect*. 2017;125:23–29.
83. Wang Y, Kloog I, Coull BA, Kosheleva A, Zanobetti A, Schwartz JD. Estimating causal effects of long-term PM2.5 exposure on mortality in New Jersey. *Environ Health Perspect*. 2016;124:1182–1188.
84. Schwartz JD, Wang Y, Kloog I, Yitshak-Sade M, Dominici F, Zanobetti A. Estimating the effects of PM2.5 on life expectancy using causal modeling methods. *Environ Health Perspect*. 2018;126:127002.
85. Goldman GT, Dominici F. Don't abandon evidence and process on air pollution policy. *Science*. 2019;363:1398–1400.
86. Cox LAT Jr. Do causal concentration-response functions exist? A critical review of associational and causal relations between fine particulate matter and mortality. *Crit Rev Toxicol*. 2017;47:603–631.
87. Cox LA Jr. Modernizing the Bradford Hill criteria for assessing causal relationships in observational data. *Crit Rev Toxicol*. 2018;48:682–712.
88. Clean Air Act, as amended by Pub. L. No. 101–159. 42 USC 7408, 1990.
89. Clean Air Act, as amended by Pub. L. No. 101–159. 42 USC 7409, 1990.